

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 23, 2002, 23:38:05 ; Search time 10173.5 Seconds

(without alignments)
5383.678 Million cell updates/sec

Title: US-09-171-553B-1

Perfect score: 3320

Sequence: 1 gaattcgcgccgcgcgcgcac.....aaaaaaaaaaaaaaaa 3320

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1472140 seqs, 8248589755 residues

Word size: 0

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenBank:*

1: gb_ba:*

2: gb_hg:*

3: gb_in:*

4: gb_ov:*

5: gb_ov:*

6: gb_ov:*

7: gb_ov:*

8: gb_ov:*

9: gb_ov:*

10: gb_ov:*

11: gb_ov:*

12: gb_ov:*

13: gb_ov:*

14: gb_ov:*

15: gb_ov:*

16: gb_ov:*

17: gb_ov:*

18: gb_ov:*

19: gb_ov:*

20: gb_ov:*

21: gb_ov:*

22: gb_ov:*

23: gb_ov:*

24: gb_ov:*

25: gb_ov:*

26: gb_ov:*

27: gb_ov:*

28: gb_ov:*

29: gb_ov:*

30: gb_ov:*

31: gb_ov:*

32: gb_ov:*

33: gb_ov:*

34: gb_ov:*

35: gb_ov:*

36: gb_ov:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

8

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3320	100.0	3320	6	A66551	A66551 Sequence 1
2	2682	80.8	8209	6	A66553	A66553 Sequence 3
3	2666	80.3	8196	6	A66552	A66552 Sequence 2
4	1769	53.3	8849	14	PEN133817	AJ133817 Porcine e
5	1680	50.6	7808	14	PEN133816	AJ133816 Porcine e
6	1659	50.0	8750	14	PEN133815	AJ133815 Porcine e
7	1637	49.3	8750	14	PEN133814	AJ133814 Porcine e
8	1606	48.4	8918	14	PEN283656	AJ283656 Porcine e
9	1091	32.9	7333	6	AF038601	AF038601 Sus scrofa
10	1091	32.9	7333	6	AF130474	AF130474 Sequence
11	848	25.5	4402	6	AX052635	AX052635 Sequence
12	797	24.0	6076	6	AX052636	AX052636 Sequence
13	797	24.0	7362	6	AX052634	AX052634 Sequence
14	786	23.7	2462	6	AX002802	AX002802 Sequence
15	786	23.7	2462	6	PERENV1	PERENV1
16	746	22.5	6076	6	AX052647	AX052647 Sequence
17	704	21.2	8132	6	AR130475	AR130475 Sequence
18	570	17.2	4918	6	AX052637	AX052637 Sequence
19	557	16.8	8132	6	AF038600	AF038600 Sus scrofa
20	547	16.5	7873	6	AX052638	AX052638 Sequence
21	547	16.5	7873	6	AX052633	AX052633 Sequence
22	474	14.3	1980	14	AF130444	AF130444 Porcine e
23	474	14.3	1980	14	PEN288585	PEN288585
24	438	13.2	2009	14	AF130473	AF130473 Sequence
25	375	11.3	8060	6	AF038599	AF038599 Sus scrofa
26	375	11.3	8132	6	AX002804	AX002804 Sequence
27	321	9.7	3482	6	PERENV2	PERENV2
28	321	9.7	3482	14	AF163267	AF163267 Porcine e
29	285	8.6	716	14	SS077599	SS077599
30	208	6.3	927	4	AF163269	AF163269 Porcine e
31	170	5.1	942	14	MLVKT10	MLVKT10
32	166	5.0	166	14	AF138386	AF138386 Porcine e
33	166	5.0	166	14	AF138389	AF138389 Porcine e
34	161	4.8	810	14	AF033259	AF033259 Sus scrofa
35	158	4.8	522	14	AF033260	AF033260 Porcine e
36	158	4.8	927	4	SS077600	SS077600
37	119	3.6	615	14	AF163269	AF163269 Porcine e
38	97	2.9	1996	10	PEN288584	PEN288584
39	65	2.0	188989	10	AC074224	AC074224 Mus muscu
40	46	1.4	176506	2	AC079956	AC079956 Mus muscu
41	46	1.4	187393	2	AC084325	AC084325 Mus muscu
42	46	1.4	195423	2	AC087151	AC087151 Mus muscu
43	45	1.4	942	4	AF274705	AF274705 Sus scrofa
44	43	1.3	192465	2	AC091333	AC091333 Mus muscu
45	41	1.2	1591	10	AF161307	AF161307 Mus caroli

ALIGNMENTS

RESULT 1

A66551

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

porcine endogenous retrovirus.

porcine endogenous retrovirus.

Viruses: Retroviral viruses: Retroviridae; Mammalian type C

retroviruses: 1-Mammalian type C virus group.

REFERENCE

AUTHORS

Galbraith,D.N., Haworth,C., Lees,G.M. and Smith,K.T.

TITLE

PORCINE RETROVIRUS

JOURNAL

Patent: WO 9740167-A 1 30-OCT-1997;

Q ONE BIOTECH LTD (GB)

FEATURES

source

1..3320

/organism="porcine endogenous retrovirus"

/db_xref="taxon:61673"

Db 2041 TCAACGGTGGCTAGAGAAATCTGAGAAATTTTCCACATTTTGGAAATACCTAAG 2100
 QY 2101 gtaataaggtcaacaatggtccagcttctgctccaggttaagtcagagagcggcgaag 2160
 Db 2101 GTAAATAGGTCAGACAAATGTCACCTTCTGTCGCCAGGTAAGTACGAGGATGCGCAAG 2160
 QY 2161 atattggattgattggaactgcatctgcatcacagaccccaagagcaggaagta 2220
 Db 2161 ATATTGGGATTGATTGGAAACCTCATTGTGCATACAGACCCCAAGCTCAGACAGGTA 2220
 QY 2221 gagaggaatgaataagacatcaagaagaccccttaaccacacagagagctgcaat 2280
 Db 2221 GAGAGATGAATAGAACATTAAAGACACCTTACCAAAATTAACACAGAGCTGCGCAT 2280
 QY 2281 aatgattgattggtctccctgctgctgctgctgcttcttaaggtaggaacaccccgagac 2340
 Db 2281 AATGATTGGATGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
 QY 2341 ttggagctgaccccttaagattgctctacaggggagaccccccgttgacagaaattgac 2400
 Db 2341 TTTGGGCTGACCCCTTGAATTTGCTTACGGGGGACCCCGCTTGCGCAGAAATTTGCC 2400
 QY 2401 ttgcacatagttctatctgctgctctccagccttctctcttaagctcagagcgctc 2460
 Db 2401 TTTGCACATAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
 QY 2461 gaattgagtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2520
 Db 2461 GAGTGGGTGAGGACGACGAGCGGTGAGAGAGCTCCGGGAGGCTTACCTACAGAGAGAGCT 2520
 QY 2521 caagttccacatcgcttccagcttccagcttccagcttccagcttccagcttccagcttcc 2580
 Db 2521 CAAGTTCCACATCGCTTCCAAAGTTGAGATTGAGTCTATGTTAGACGCCACCGTCCAGGA 2580
 QY 2581 aacctcagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2640
 Db 2581 AACCTCGAGACTCGGTGAGAGGACCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
 QY 2641 aagtcagaagaatcccaactgcatgcatgcatgcatgcatgcatgcatgcatgcatgcatgcat 2700
 Db 2641 AAATCGAAGGAATCCCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2700
 QY 2701 gattcgggtggaagacggaagaagcagcagcagcagcagcagcagcagcagcagcagcagc 2760
 Db 2701 GATTGGGGTGAAGACCCCAAAAGACTGAATAATCCCTTAAAGCTTCCCTCCATCCCGT 2760
 QY 2761 gtctctactctgctcaataactcctcaagtaactgtaaaagccttgtagagagccgga 2820
 Db 2761 GTTCTTACTCTGCTCAATTAATCTCCCAAGTTAATGTAAGCCCTTGTGAGACACCCCGAA 2820
 QY 2821 ctcccaataacccttactctcactgctgctgctgctgctgctgctgctgctgctgctgctgct 2880
 Db 2821 CTCCCAATAACCTTATCTCTACCTGCTTACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880
 QY 2881 taacagcactcaagggaggggagccttggggagcctggtgagcctggaattatgctgct 2940
 Db 2881 TAACAGCACTCAAGGGAGGCTCTTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2940
 QY 2941 tgcagcaataactcctgctcctcaatgacagggacacacacacacacacacacacacacacac 3000
 Db 2941 TGCATCAATTAATCTCTGCTGCTCAATGACAGCCACACCCCGATGACTGCTGCTGCTGCT 3000
 QY 3001 cgggttctacgttgcacagggagggagggagggagggagggagggagggagggagggagggag 3060
 Db 3001 CGGTTTACGTTTGCAGGAGACCCCAAAATTAATGAATAATTTGTGAATCTCTCAGGA 3060
 QY 3061 ttctccttgcagcaatgagcgtcgtcaataactcctcaatgagggagggagggagggagggag 3120
 Db 3061 TTTCTCTTGCAGCAATGAGGCTGCTGCTTAACTTAATGATGGGATTTGGAATGGGCACT 3120
 QY 3121 ctctcagcaagcagagtaagttactcttctgtaacaacacacacacacacacacacacacacac 3180
 Db 3121 CTCTCAGCAAGCAGAGTAAGTTACTCTTTTGTTAACAATCTCAACAGTTAATATCAAT 3180

QY 3181 taatagccatgaggaatggaagatggaacagcgggtacacaaagatgtagcga 3240
 Db 3181 TAATTATGGCCATGGGAGATGGAAGATGGAACAGCGGTCACAAAAGATGTCGAAA 3240
 QY 3241 taagcaatagcctgctcctgctgctgctgctgctgctgctgctgctgctgctgctgct 3300
 Db 3241 TAAGCAATAATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300
 QY 3301 aaaaaaaaggg 3360
 Db 3301 AAAAAAAAAAAAAAAAAA 3320

RESULT 2
 LOCUS A66553 8209 bp DNA
 DEFINITION Sequence 3 from Patent WO9740167.
 ACCESSION A66553
 VERSION A66553.1 GI:4538106
 KEYWORDS
 SOURCE
 ORGANISM
 porcine endogenous retrovirus.
 porcine endogenous retrovirus.
 Viruses; Retroviral viruses; Retroviridae; Mammalian type C
 1 (bases 1 to 8209)
 Galbraith, D. N., Hawthorth, C., Lees, G. M. and Smith, K. T.
 PATENT: WO 9740167-A 3 30-OCT-1997;
 Q ONE BIOTECH LTD (GB)
 Location/Qualifiers
 1. 8209
 /organism="porcine endogenous retrovirus"
 /db_xref="taxon:61673"
 588..2162
 /codon_start=1
 /product="VIRAL CORE PROTEIN"
 /protein_id="CAB39347.1"
 /db_xref="GI:4538107"
 /translation="MGQIVTTPILSLIDHWTEVRSNHNLSVYKKGPMQFCASEMP
 TEDVWPESEGFENSEITLAKALITFTGDSHPDOPEYILTMODLADPPMYPWMLN
 KPRKGPRIILAKRNKHSAREKPESSYLPRDRAADLAGPTCSPTISSTGCCG
 TSAPGAPVYEGPAGRSRGAATPERTEDEILPLRTYGPMPGGLQIPQVPESS
 ADLYNKTINHPESSEDFQRLIGVESLIMSHPMDDCQOLLOTLFTTEERILLTA
 RNVPAGDGRPTOLONEIDMGFELTPEGMVYTAGRSLSLITQALVAGIRGASRP
 TNLAKREVMQGNPEPSEVLEIMAPRPPTPTSEAKASVALAFTGOSALDIR
 KTIQRLGLOEALRDVREAKVYRRETEBEKORERERERERERERERERERERER
 KTIQAVVEGKSSRERDRPKIRSGPRSGNLGNTP.LDKDOCAVCEKGMARNCPK
 <2163..5747
 /codon_start=1
 /product="REVERSE TRANSCRIPTASE"
 /protein_id="CAB39348.1"
 /db_xref="GI:4538108"
 /translation="GRGSDPLPEPRVITKVEQPPPELVDTGAEHSVLLQPLGKIKE
 KSWGAGATGQROYPMTTRRTVDGCGVRVTHSLVPECEVPLIGRLDITKGAQISF
 EGREPEVSNKPLITVLTOIDPEYRLISQVAPDDQISWLEQFOAAETVAGAGLA
 KOVPOVIOKASATPVSVROYRPLSRARSGIMPVQVRLIOGILVPOVSNMPILP
 VRRPGTNDYRPVODLREYNKRVODIHPTVNPYNLSALPPEPRMNTYLDIKDAFTCL
 RLHPTSOPLFAFEMRDRGTQTLMTLRFQGFKNSTPITFEALHRLANRLOHPO
 VTILOYVDLLLAGATKQDCELEKALLETLSLQSGRASKAKQALCREVLYLSLGR
 GQORULTBARKKTVQVLPAPITAKQVREFGTAGRCRILNBSAGREIKNKEETISLE
 GFSNAPBEHOKFADAIRKALISAPALAPVTPKPTPLTYDERGAVRGVLTQTPMR
 PVALTSKRLDPVAGSMVCLKATAVAAILVKADKTLTIGONITVIAPALENTVROP
 DRWMTNARMTHYOSILLETTEVTPAPPAALNPVTLPEETDEPVHDOHLEIETGR
 KDLDIDPLTEGEVLTWPTDGSYVVEKRRAGAAVVDGRTIWAASLEPGSAKTEIM
 ALTDALAEGRKSTINTYDSRYAPRAHAHGAIVKQRLISAGREIKNKEETISLE
 YLLEDEKELIITHGQHKOKDLISGONOMADRVAKQAQAVNLPIIETPKAREPPO
 VRTSPYHVLRLTEVADSVVKKHCVPOLVANAPSLIPGKRLRSGHGAHVEDVTEVK
 PARYKAKTILVEVDVDFSGVWEAVPTPKETSVGNKTIIEELPRFGIPVYVSDNGPA
 FVAOVSGAKLILGIDMKHCAVROSSGQVVERNNPTKETLRLTTEGINDMALL
 PVLFRVNTPGQFLTEYELLGPPPLAIAIARHSAVDVLSQPLSRKALEWRO

CDS

QY 3181 taatagccatgaggaatggaagatggaacagcgggtacacaaagatgtagcga 3240
 Db 3181 TAATTATGGCCATGGGAGATGGAAGATGGAACAGCGGTCACAAAAGATGTCGAAA 3240
 QY 3241 taagcaatagcctgctcctgctgctgctgctgctgctgctgctgctgctgctgctgct 3300
 Db 3241 TAAGCAATAATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300
 QY 3301 aaaaaaaaggg 3360
 Db 3301 AAAAAAAAAAAAAAAAAA 3320

RAMKOLREAVSGDLOVPHRFOVDSYVVRHAGNLETRMKPVLVLTPTFAVKE
 GIPWIIHSHVKKPAPPDPSOMRAKTEMLKRLHRLVPISSNNNSPGO"
 /codon.start-1
 /product="VIRAL ENVELOPE PROTEIN"
 /protein_id="CA39349.1"
 /db_xref="GI:4538109"
 /translation="MNPRLSPRLPTRGEPKRLIPLSFASIAMFLTLTTTQASK
 RLIDNSPHRPLSLIILIDPTGTVNSTRGVARGTWPELHFCRLINPAKSTP
 PNIVSYGFCCPEGEKTKCGSGSPCRMSCVTSNDGDMKPISLIDVWFSEVY
 SGPEYKMKMLIKDKSCSPSDLDYLTISTEPKTKSKVKNYELGNSFLVGGAG
 STILRIETIEGEPVAMGPKVLEQEPALPEPHNIPVPLSLPDTTQPSNS
 TTGLIPTNPENSBGVKTKGRLFSLIGARALNSDPAOTSLSSGPYE
 GKAKREKNVTRKRNOCITWGSRNKLITLSEVSGTICAKAPRSHOHCYETVYEA
 SENOTIYGVYRNMWACNTGLPCVSTSVENOSKDFVVOIVPRYIYPREVILDEYD
 YRINPKREPISTLAVMLIGTAVGVTGTAIITLPQOLGJELHMAETDLRA
 LKESVSNLEESTISLSEVYIIONRGLDLPLRGGGLCAALKECECPYDHSAITDSM
 NKLRKLERRRERERADOGPEGEWNSPMWTLTSLALTGPLVILLTLTVGPECLINR
 FVAVERVSAVOIWLKQOYQGLISGENDL"
 FVAVERVSAVOIWLKQOYQGLISGENDL"
 BASE COUNT 2168 a 2064 c 2154 g 1823 t
 ORIGIN
 Query Match 80.8%; Score 2682; DB 6; Length 8209;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2682; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 21 agatgctctctctgctgagatlaaaccactaaccacttttgcctcgaatg 80
 2972 agatgctctctctgctgagatlaaaccactaaccacttttgcctcgaatg 3031
 81 gagagatccaggtacggaagaacccgagctcactgagaccgagctcccaagggt 140
 3032 gagagatccaggtacggaagaacccgagctcactgagaccgagctcccaagggt 3091
 141 caagaactccgagacatcttgaagaagccctacacagagactgagcccaacttcaagt 200
 3092 caagaactccgagacatcttgaagaagccctacacagagactgagcccaacttcaagt 3151
 201 ccaacacccctcaggtacacctctccagtaagtgagactgcttctgaggaagcaac 260
 3152 ccaacacccctcaggtacacctctccagtaagtgagactgcttctgaggaagcaac 3211
 261 caacaggaactgctgagaggtacagagcactactgctggaatgtctgactagctga 320
 3212 caacaggaactgctgagaggtacagagcactactgctggaatgtctgactagctga 3271
 321 cagagacctctgtaagaagcccaagattgcaagaagaggttaacatacttggggtacag 380
 3272 cagagacctctgtaagaagcccaagattgcaagaagaggttaacatacttggggtacag 3331
 381 ttgagggggcagagcagatgagtgagagagcagcaggaagaaactgtatgctcagatcc 440
 3332 ttgagggggcagagcagatgagtgagagagcagcaggaagaaactgtatgctcagatcc 3391
 441 gggcccaacacagcaagaagtgagagattttggggaagcgtggattttgcaagact 500
 3392 gggcccaacacagcaagaagtgagagattttggggaagcgtggattttgcaagact 3451
 501 gtagatccgggggtttgagacctgagagcccaactaaccgcttaaccaagaagaagg 560
 3452 gtagatccgggggtttgagacctgagagcccaactaaccgcttaaccaagaagaagg 3511
 561 gtagatccgggggtttgagacctgagagcccaactaaccgcttaaccaagaagaagg 620
 3512 gtagatccgggggtttgagacctgagagcccaactaaccgcttaaccaagaagaagg 3571
 621 gtagatccgggggtttgagacctgagagcccaactaaccgcttaaccaagaagaagg 680
 3572 gtagatccgggggtttgagacctgagagcccaactaaccgcttaaccaagaagaagg 3631
 681 gtagatccgggggtttgagacctgagagcccaactaaccgcttaaccaagaagaagg 740

3632 GCGTAAGGAGTAGCCGAGAGCTTTTAAACCAACCAACCTAGACATGAGAGACCTGT 3691
 741 tgcctaccctgcaagaagagcttgcctgtagcagagtggtggccglatgctgaaggc 800
 3692 tgcctaccctgcaagaagagcttgcctgtagcagagtggtggccglatgctgaaggc 3751
 801 taccgagcgttgagcactggtgcaagaagcctgagaaatgactttgggaagaaat 860
 3752 taccgagcgttgagcactggtgcaagaagcctgagaaatgactttgggaagaaat 3811
 861 aactgtaataagcccccaatgcatgtagaagaacatgcttgcgagccccaagacagatg 920
 3812 aactgtaataagcccccaatgcatgtagaagaacatgcttgcgagccccaagacagatg 3871
 921 gacacagcccgagacagac 980
 3872 gacacagcccgagacagac 3931
 981 tccacacagcccgagac 1040
 3932 tccacacagcccgagac 3991
 1041 tccatgattgcatcaactatgattgagagagctggggtccggaagacactaagacat 1100
 3992 tccatgattgcatcaactatgattgagagagctggggtccggaagacactaagacat 4051
 1101 accgctgagctgagagagctgtagcactgctgctgagagagagctgtagcactgtag 1160
 4052 accgctgagctgagagagctgtagcactgctgctgagagagagctgtagcactgtag 4111
 1161 taagaaggaatgctgagagagctgtagcactgctgctgagagagagctgtagcactgtag 1220
 4112 taagaaggaatgctgagagagctgtagcactgctgctgagagagagctgtagcactgtag 4171
 1221 ggcggaagagagctgtagcactgctgctgagagagagctgtagcactgtagcactgtag 1280
 4172 ggcggaagagagctgtagcactgctgctgagagagagctgtagcactgtagcactgtag 4231
 1281 ggcggaagagagagctgtagcactgctgctgagagagagctgtagcactgtagcactgtag 1340
 4232 ggcggaagagagagagctgtagcactgctgctgagagagagctgtagcactgtagcactgtag 4291
 1341 ggcggaagagagagagctgtagcactgctgctgagagagagctgtagcactgtagcactgtag 1400
 4292 ggcggaagagagagagctgtagcactgctgctgagagagagctgtagcactgtagcactgtag 4351
 1401 ggcggaagagagagagctgtagcactgctgctgagagagagctgtagcactgtagcactgtag 1460
 4352 ggcggaagagagagagctgtagcactgctgctgagagagagctgtagcactgtagcactgtag 4411
 1461 ggcggaagagagagagctgtagcactgctgctgagagagagctgtagcactgtagcactgtag 1520
 4412 ggcggaagagagagagctgtagcactgctgctgagagagagctgtagcactgtagcactgtag 4471
 1521 ggcggaagagagagagctgtagcactgctgctgagagagagctgtagcactgtagcactgtag 1580
 4472 ggcggaagagagagagctgtagcactgctgctgagagagagctgtagcactgtagcactgtag 4531
 1581 ggcggaagagagagagctgtagcactgctgctgagagagagctgtagcactgtagcactgtag 1640
 4532 ggcggaagagagagagctgtagcactgctgctgagagagagctgtagcactgtagcactgtag 4591
 1641 ggcggaagagagagagctgtagcactgctgctgagagagagctgtagcactgtagcactgtag 1700
 4592 ggcggaagagagagagctgtagcactgctgctgagagagagctgtagcactgtagcactgtag 4651
 1701 ggcggaagagagagagctgtagcactgctgctgagagagagctgtagcactgtagcactgtag 1760
 4652 ggcggaagagagagagctgtagcactgctgctgagagagagctgtagcactgtagcactgtag 4711
 1761 ggcggaagagagagagctgtagcactgctgctgagagagagctgtagcactgtagcactgtag 1820
 4712 ggcggaagagagagagctgtagcactgctgctgagagagagctgtagcactgtagcactgtag 4771

```

QY 1821 ggcctgactcgggtggtcaaacatgtgtccctgcagctggttaactgtaactcctcag 1880
    |||
Db 4772 ggctgactcgggtggtcaaacatgtgtccctgcagctggttaactgtaactcctcag 4831
    |||
QY 1881 aatactccaggaagaagactaagggaagccacccagcgctcactggaagtgact 1940
    |||
Db 4832 aatactccaggaagaagactaagggaagccacccagcgctcactggaagtgact 4891
    |||
QY 1941 cactgaggttaaacccggtcgaataacgaataatattgtgttttttgtagaacctt 2000
    |||
Db 4892 cactgaggttaaacccggtcgaataacgaataatattgtgttttttgtagaacctt 4951
    |||
QY 2001 ttcaagatggttagaggtcttactactaagaagaagacttcaacgtggtggttaagaa 2060
    |||
Db 4952 ttcaagatggttagaggtcttactactaagaagaagacttcaacgtggtggttaagaa 5011
    |||
QY 2061 aatactggaagaaatttttccaaagatttggaaatcactaaggttaaggttcaagaatg 2120
    |||
Db 5012 aatactggaagaaatttttccaaagatttggaaatcactaaggttaaggttcaagaatg 5071
    |||
QY 2121 tccagcttcgttgcacgaaglaaglcagagactggtcgaagatttggttgatgaa 2180
    |||
Db 5072 tccagcttcgttgcacgaaglaaglcagagactggtcgaagatttggttgatgaa 5131
    |||
QY 2181 actgacttgcatacagaaccccaagctcaggaaggttagaagaa tgaataaacat 2240
    |||
Db 5132 actgacttgcatacagaaccccaagctcaggaaggttagaagaa tgaataaacat 5191
    |||
QY 2241 taaagagaccttaccgaattgacacagagactggtcattatgattggtgctcct 2300
    |||
Db 5192 taaagagaccttaccgaattgacacagagactggtcattatgattggtgctcct 5251
    |||
QY 2301 gcccttctgcttcttaaggtgtagaagaaacccctggaagcttgggttgacccctatga 2360
    |||
Db 5252 gcccttctgcttcttaaggtgtagaagaaacccctggaagcttgggttgacccctatga 5311
    |||
QY 2361 attgcttcaagggggaacccccccgttggcaagaattgcttgcacatgctgattg 2420
    |||
Db 5312 attgcttcaagggggaacccccccgttggcaagaattgcttgcacatgctgattg 5371
    |||
QY 2421 gctgcttccagccttctgcttcttaaggtgtagaagaaacccctggaagcttgggttgacccctatga 2480
    |||
Db 5372 gctgcttccagccttctgcttcttaaggtgtagaagaaacccctggaagcttgggttgacccctatga 5431
    |||
QY 2481 gtagaagcagctccgggagagcctactcaggaagagacttgcagttccacatgccttca 2540
    |||
Db 5432 gtagaagcagctccgggagagcctactcaggaagagacttgcagttccacatgccttca 5491
    |||
QY 2541 agtggagatcagctcattatgttgaagcgaacgttgcaggaacccctgcagactcgttggaa 2600
    |||
Db 5492 agtggagatcagctcattatgttgaagcgaacgttgcaggaacccctgcagactcgttggaa 5551
    |||
QY 2601 ggaacctatctgtaactttagacacacacagcgtgtgaaagtcgaagaatctccac 2660
    |||
Db 5552 ggaacctatctgtaactttagacacacacagcgtgtgaaagtcgaagaatctccac 5611
    |||
QY 2661 ctgagatcagtcaccccaaggttaagcgcgccacactccga 2702
    |||
Db 5612 ctgagatcagtcaccccaaggttaagcgcgccacactccga 5653
    |||

```

```

RESULT 3
LOCUS A66552 8196 bp DNA PAT 29-MAR-1999
DEFINITION Sequence 2 from Patent WO9740167.
ACCESSION A66552
VERSION A66552.1 GI:4538105
KEYWORDS
SOURCE
ORGANISM
porcine endogenous retrovirus.
Virus; Retroviral virus; Retroviridae; Mammalian type C
retroviruses; I-Mammalian type C virus group.

```

```

REFERENCE 1 (bases 1 to 8196)
AUTHORS Galbraith,D.N., Haworth,C., Lees,G.M. and Smith,K.T.
TITLE PORCINE RETROVIRUS
JOURNAL Patent: WO 9740167-A 2 30-OCT-1997;
Q ONE BIOTECH LTD (GB)
FEATURES
source location/Qualifiers
1..8196
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
BASE COUNT 2165 a 2061 c 2146 g 1820 t 4 others
ORIGIN
Query Match 80.3%; Score 2666; DB 6; Length 8196;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 21 agatgctctctctgctgagatctacacccactagcaaccatttgcctcgaatg 80
    |||
Db 2958 agatgctctctctgctgagatctacacccactagcaaccatttgcctcgaatg 3017
    |||
QY 81 gtagatccaggtacgggaagaaacccggcagctcacttgcagccagctgcccagaggt 140
    |||
Db 3018 gtagatccaggtacgggaagaaacccggcagctcacttgcagccagctgcccagaggt 3077
    |||
QY 141 caagaactcccgaccactcttgaagaaagccctacacagggacttgcgaactcagat 200
    |||
Db 3078 caagaactcccgaccactcttgaagaaagccctacacagggacttgcgaactcagat 3137
    |||
QY 201 ccaaacacctcaggtgacccctcctcagtaagatgacgttcttgcggagagccac 260
    |||
Db 3138 ccaaacacctcaggtgacccctcctcagtaagatgacgttcttgcggagagccac 3197
    |||
QY 261 caaagaagactgcttagaaggttaagaaagcaactcgtggaattgctcagtaagct 320
    |||
Db 3198 caaagaagactgcttagaaggttaagaaagcaactcgtggaattgctcagtaagct 3257
    |||
QY 321 cagaagcctctgctgaagagccagatttgaagaaagaggttaacatcttggtgacg 380
    |||
Db 3258 cagaagcctctgctgaagagccagatttgaagaaagaggttaacatcttggtgacg 3317
    |||
QY 381 ttgctggggcgagcagcgtatgctgacgaggaagcaagaaagaaactgactcagatacc 440
    |||
Db 3318 ttgctggggcgagcagcgtatgctgacgaggaagcaagaaagaaactgactcagatacc 3377
    |||
QY 441 ggcctccacacacagcacaagaagtggagaggttcttggggagcagcttgcagact 500
    |||
Db 3378 ggcctccacacacagcacaagaagtggagaggttcttggggagcagcttgcagact 3437
    |||
QY 501 gtagatcccgaggttgcagacttagcagcccaacttaccgcttaacccaagaagaag 560
    |||
Db 3438 gtagatcccgaggttgcagacttagcagcccaacttaccgcttaacccaagaagaag 3497
    |||
QY 561 ggaattctcctggctccttgaagcaaggaagcatttgaatgctatcaaaaagccctgtc 620
    |||
Db 3498 ggaattctcctggctccttgaagcaaggaagcatttgaatgctatcaaaaagccctgtc 3557
    |||
QY 621 gagcgacacctgcttgcctcctcctgaagcgaataacaccttacccttattgtgata 680
    |||
Db 3558 gagcgacacctgcttgcctcctcctgaagcgaataacaccttacccttattgtgata 3617
    |||
QY 681 gcgtaaaggaagtagcccgagagattttaaaccacacccctagagcaatgaagagactgt 740
    |||
Db 3618 gcgtaaaggaagtagcccgagagattttaaaccacacccctagagcaatgaagagactgt 3677
    |||
QY 741 tgcttaccttcaaaagagacttgaatcctgtagcagtggttgcgcctatgtcttaagagc 800
    |||
Db 3678 tgcttaccttcaaaagagacttgaatcctgtagcagtggttgcgcctatgtcttaagagc 3737
    |||
QY 801 tatcgacgctgtgccaactatggttcaagaaagcgttgaacatgtgacttggagcaaatat 860
    |||
Db 3738 tatcgacgctgtgccaactatggttcaagaaagcgttgaacatgtgacttggagcaaatat 3797
    |||

```

QY 861 aactgtaataagcccccatgcatcttgagaaacatgctgcgagcccccaagcatgat 920
 DB 3798 AACTGTAATAAGCCCCCATGCACTTGAGAAACATCGTTGGAGGCCCCAGACCATGAT 3857
 QY 921 gaacaaagcccgatgaaccatcatcaaaagcctgtcttcacagaagagtcactctgc 980
 DB 3858 GAAACAAAGCCCGATGACCCACATCAAAAGCCTGCTTCACAGAGAGGCTCATCTTCCG 3917
 QY 981 tccacagccgcctctaaccttgcacactcttctgcctgaagaagctgtatgaacagtgac 1040
 DB 3918 TCCACAGCCGCGCTCTCAACCCCTGCACTCTTGCCTGAAGAGACTGTGAACCACTGAC 3977
 QY 1041 tcatgtttgcatcaactatgatgatgaagaactctgggttcgcaagagcccttaacagat 1100
 DB 3978 TCATGATTCCTCAATCAACTATTGATTAGAGACATGGGGTCCCAAGACCTTACAGACAT 4037
 QY 1101 accgctgactgtgagaagctgtaaacctgttctacatgacggaagcagctatgtgttgaaag 1160
 DB 4038 ACCGCTGACTGAGAGAGTGTCAACCTGTTCACTGACGGAACAGCTATGTGTTGAAG 4097
 QY 1161 taagaagatgctcgagcgagctgtgttgagacggagcccgacagctgtgccaagagct 1220
 DB 4098 TAAGAGATGCTGGGGCGCACTGTGTGAGACGGACCCGACGATCTGGGCCAGACGCT 4157
 QY 1221 gccggaaggaacttcagcgcaaaagctgagctcatgagccctcaagcagcttgcgagct 1280
 DB 4158 GCCGGAAGGAACCTTCACGCCAAAGCTGACCTCATGSCCTCAGCAAGCTTTGCGGCT 4217
 QY 1281 gccggaaggaatccataaacaattatacgaagcagagctatgcttgcctgtgacatgaca 1340
 DB 4218 GGCGGAAGGGAATCCATAACATTTATACGACACAGATGTGCTTGGCACATGACACA 4277
 QY 1341 cgtacaggggcatctatatacaaaaggggttgcttactctcagcagaggaagaaataa 1400
 DB 4278 CGTACAGGGGCACTATATAACAAGGGGTTGCTTACTCTACAGAGGAGGAAATATAA 4337
 QY 1401 gaacaaaggaatcttaagcctatagaagccttactatctgcccacaaagctagctat 1460
 DB 4338 GAACAAAGGAATTTTAAGCCTATTAGAAGCCTTACATTTGCCAAAAAGCTAGCTAT 4397
 QY 1461 tatacactgtcttgagacatagaagaagcacaagctcctatctatagaaggaacagatgac 1520
 DB 4398 TATACACTGTCTTGACATCAGAAAGCAAAATCTCATATGAGGGAGACCAAGTGAC 4457
 QY 1521 tgacaggggttgcaagcagcagccagcagctgtttaaactctgcctcataatagaagcgc 1580
 DB 4458 TGACCGGGTTCACAGCAGGACGCCAGGCTTTAACTCTTGCCATATATAGAAACGCG 4517
 QY 1581 caaagccccaagaccagacagacagtaacacctagaagactgtgcaagagaataaagat 1640
 DB 4518 CAAAGCCCAAGACCCAGACGACAGTACACCTTAGAAAGCTGGCAAGATATAAAGAT 4577
 QY 1641 agaccagttctctgagactccgagagggagactgtctactactatagaggaagaaatgct 1700
 DB 4578 AGACCAAGTTCTTGAGACTCCGGAGGGACCTCTATACCTCATATAGGGAAGAAATCTT 4637
 QY 1701 gccccaacaagaaggtctagaataatgccaacagatatacgtcttaacccaactagaagac 1760
 DB 4638 GCCCCACAAGAAGGTTAGATATGTGCCAAAGATACATCGTCTAACCCACCTAGAGAC 4697
 QY 1761 taacaacctgacagattgttcagaaacatccctctatcatgtctctgagctacagagat 1820
 DB 4698 TAAACACCTTCACAGATTTGTTCAGAACATCCCTTATCATGTCTGAGGCTACAGAGAT 4757
 QY 1821 ggcgtgactcggtgtcaaaccttctgtccctccagctggttaagtcaatccctcag 1880
 DB 4758 GGCCTGACTCGGTGTCAACATTTGTGCTTCCCTGCACTGTTATATCTTATCTTCCAG 4817
 QY 1881 aatacctcagaagaagactaaagggaagccacccagcgctcactggaagcttgagctt 1940
 DB 4818 AATACCTCCAGAAAGAGACTAAGGGAAGCACCAGGCGCTACACTGGGAATGAGACTT 4877
 QY 1941 cactgagtgtaagccgcttaataatcggaacaataatctatgtgttttctgtaacacct 2000

DB 4878 CACTGAGTAAGCCCGGCTAAATACGGAACAAATATCTATTGGTTTGTGACACCTT 4937
 QY 2001 ttcaagatggtgaagagctatctcactaaagaagaactcaacgctgtgtgtgaagaa 2060
 DB 4938 TTCAAGATGGTAGAGGCTTATCTACTACTAGAAAGACCTTCACCCGCTGGCTAAGAA 4997
 QY 2061 aatacctgagaatatttccaaagattggaatacctaagglaaaggttcagaactgt 2120
 DB 4998 AATACCTGAGGAATATTTTCCAAAGATTGGAATACCTTAAGGTAAATAGGCTCAGCAATGG 5057
 QY 2121 tccagcttctgttgcacagtaagtcagagctgtgccaagatttggagattgattgaa 2180
 DB 5058 TCCAGCTTCTGTTGCCACAGTAGTACAGGACTGGCCAAAGATTTGGGGATTGTTGAA 5117
 QY 2181 actgcatctgtacatacagccccaagctcgaagaagctagaagagatgaataagaaat 2240
 DB 5118 ACTGCTTTGTGCTACAGACCCCAAAAGCTCAGACAGAGTAGAGATGATAGACCAT 5177
 QY 2241 taagaagaccttaccacatctgacacagagactgtgcatatgaattgtgtgtctcct 2300
 DB 5178 TAAAGAGACCTTACCAAAATGACACAGAGACTGGCATTAATGATTTGATGGCTCTCT 5237
 QY 2301 gcccttgtgtctttaaaggtgaagaaaccccttgagacgtttgtgtgtgtgtgtgtgt 2360
 DB 5238 GCCCTTGTGCTTTTAAAGGTGAGGAAACCCCTGACAGATTTGGCTGACCCCTATGA 5297
 QY 2361 attgtctacggggagcccccccgctgtgcagaatattgcttgcataatgtgtatgt 2420
 DB 5298 ATTGCTCTACGGGGAGACCCCCCGTGGACAAATTTGCTTGCACATPATGTGATGT 5357
 QY 2421 gctgtcttccagccttgttctcctagctcagaagcgtcgtgtgtgtgtgtgtgtgtgt 2480
 DB 5358 GCTGCTTTCCACAGCTTTTGTGCTCTAGCTCAAGGCGCTCGATGGTGTGAGCACGAGC 5417
 QY 2481 gtggaagcaactccgggaagcctcactcaggaagaagactgtcaagcttccacatgttcca 2540
 DB 5418 GTGGAAGCACTCCGGGAGGCTCTACTCAGGAGAGACTTGCAGATTTCCACATCGCTTCCA 5477
 QY 2541 agttgagatcagctatgtltagacgcacagctgtcagaagaacccctcgaagctcgtgtgaa 2600
 DB 5478 AGTTGAGATTTCAAGTATTTAGAGCGCACCGCTGACGAGAAACCTGACACTCGGTGAA 5537
 QY 2601 ggaacctatctgtacttttgcacacaccaaaggtgtgaaagtcgaagatccctcac 2660
 DB 5538 GGGACCTTATCTGTACTTTTGAACACACCAAGCGCTGTGAAGTGAAGGATCCAC 5597
 QY 2661 ctgagatcatgcatccacgttaagc 2686
 DB 5598 CTGATTCATGATCATCCACGTTAAGC 5623

RESULT 4
 PEN133817 8849 bp DNA VRL 07-JAN-2000
 LOCUS Porcine endogenous retrovirus type C proviral gag, pol and env
 DEFINITION Porcine endogenous retrovirus type C (class A, clone 42).
 ACCESSION AJ133817
 VERSION AJ133817.1 GI:6688947
 KEYWORDS env gene; env protein; gag gene; gag protein; pol gene; pol protein.
 SOURCE porcine endogenous retrovirus.
 ORGANISM porcine endogenous retrovirus.
 Viruses; Retroid viruses; Retroviridae; Mammalian type C retroviruses; 1-Mammalian type C virus group.
 1 (bases 1 to 8849)
 Cauderna,F., Fischer,N., Boller,K., Krach,U., Kurth,R. and Toenjes,R.R.
 TITLE Molecular characterization of Human-tropic and Replication-competent Porcine Endogenous Retroviruses
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 8849)
 AUTHORS Toenjes,R.R.

Db 4279 TATCGAGCTGTGGCATACTGCTCAAGAGACGCTGACAAATTCCTTTGGACAGATAT 4338
 Qy 861 aactgtaataagcccccatgcatcttgagaaacatcgctgcagccccagacagatgag 920
 Db 4339 AACTGTATATGCCCCCATGATGATGAGAAACATCTTCCGACAGCCCCAGACGATGAT 4398
 Qy 921 gaccacagcccgatgacacacatacaagcctgctctcacaagagaaggtcaacttgcg 980
 Db 4399 GACCAAGCGCCGATGACCCATATCAAAAGCCCTGCTTCCACAGAGAGGCTCACTTCCG 4458
 Qy 981 tccacaagcccgctcaaacctgcacctcttcgctgagagagacgatgaaacagtgag 1040
 Db 4459 TCCACAGCGCTCTCAACCTGCACTCTTCTGCTGAGAGACATGATGAACCATGTAC 4518
 Qy 1041 tcatgattgcatcaacatctgattagagagactgggtccgcaagagacttaacaacat 1100
 Db 4519 TCATGATGTGCATCACTATGATGAGAGACTGGGGTCCGACAGACCTTACACACT 4578
 Qy 1101 accgctgactggaagagtgctaaactggttcaactgacggaagacgactatggtggaag 1160
 Db 4579 ACCGCTGACTGGAAGAGTCACTACCTGCTCACTGACGACAGACGCTATGTGTGGAAG 4638
 Qy 1161 taagagatgctgggggagcagtggtgagcggagaccgacagatctggcagagacgt 1220
 Db 4639 TAAAGAGATGGCTGGGGGACGTGGTGGAGCGGACCCGACATCTGGCCACGACCT 4698
 Qy 1221 gccggaagagactcagcgcaaaaggctgagctcagctcagcctcagcgaagcttgctgct 1280
 Db 4699 GCCGGAAGAGAACTTCAAGCCCAAAAGGCTGAGCTCATGGCCCTCACCAAGCTTGGGCT 4758
 Qy 1281 ggcggaagagaaatccataacattataagagcagcaggtatgctcttgagactgca 1340
 Db 4759 GGGCGAAGGGAATTCATTAACATTTATACGACACAGGATAGCCCTTGGCGACTGCA 4818
 Qy 1341 cgtacaaggggcatctatacaaaagggtgcttcaactcagcagcgagagagaaataa 1400
 Db 4819 CGTACAGGGGCCATCTATTAACAAAGGGGCTGCTTACTCGACAGGAGGGAATTA 4878
 Qy 1401 gaacaaagagaaattcctaagcctatagaagccttaacatttgccaaaaagagctagcat 1460
 Db 4879 GAACAAAGAGAGAAATTTAGACCTTATGAGAGCTTACATTTGCCAAAAAGGCTTACTAT 4938
 Qy 1461 tataactgctcgtgagacataagaagccaaagatctcatalctagaaggagacagatgac 1520
 Db 4939 TATACACTGCTCGGATGAGATAGAAAGCCAAAGATCTCATATCTAGAGGACACAGATGG 4998
 Qy 1521 tgaccgggttgcaagcagcagcccgagctgttaaactctgctcatalaataagaacagcc 1580
 Db 4999 TGACCGGGTTGCCAAGCAGGACGCCAGGCTGTAACTCTGCTATATATAGAAAGGCC 5058
 Qy 1581 caaaagcccaagacccagacagacagtaacacccctagaagactggaagagataaaagat 1640
 Db 5059 CAAAAGCCCAAGACCCAGAGACAGTACACCTAGAAAGATGGCAAGAGATTAATAAAGAT 5118
 Qy 1641 agaccagttctcagagactccgagggagactgcttaactcactaatactggaagaaactc 1700
 Db 5119 AGACAGATTCTCTGAGACTCCGGAGGGGACTGCTACTACTCATATGSGAAGAAATCT 5178
 Qy 1701 gcccacaagaagggctagaataatgctcaacagatacatcgtctcaaccccaactaggaac 1760
 Db 5179 GCCCCACAAGAAAGGGTTAAATATGTCCAAAGATACATGCTCTACCCACTAGGAAC 5238
 Qy 1761 taacacactcagcagagttgctagaacatacccttactcattgctctgaagatacagaagt 1820
 Db 5239 TAAACACTCTGAGAGTTGGTCAAGACATCCCTTATCATAGTTCTGAGGGTACAGAGAT 5298
 Qy 1821 gctgactcgtggtcaaaatggtgctcctcagcagctggttaatgacttaactcctcag 1880
 Db 5299 GGCTGATCTGGTGTCAAACTTCTGTGCCCTGCCAGCTGTTAATGCTAATCTTCCAG 5358
 Qy 1881 aatactccagaaagagactaaagggaagcacaacagcgctcactggaagtgagact 1940
 Db 5359 AATACCTCAAGAAAGAGACTAAAGGGGAACCAACCCAGGCGCTACAGTGGAAAGTGAAT 5418

Qy 1941 cactgagglaaagcccggttaataacggaacaaatatactatgctcttctgagacact 2000
 Db 5419 CACTGAGTAAAGCCGCTTAATACGAAACAAATATCTATGCTTTTGTGAGCACCTT 5478
 Qy 2001 ttcagaatggttagaggtcttccactaagaagaagactcaacgcttgggtgagaa 2060
 Db 5479 TTCAGGAATGGGTAGAGGCTTATCTTACTTAAAGAAAGAGCTTCAACCGTGGGTGA 5538
 Qy 2061 aatactgaggaatcttccaaagattggaatacctaaggttaataaggtlcaagaatg 2120
 Db 5539 AATACTGGAGGAATTTTCCGAGATTTGGAATACCTAAGGTAAATCGGGTCAGCAATGG 5598
 Qy 2121 tccagcttgcgttgcacagtaagtcagggactggccaagatattgggatttgaa 2180
 Db 5599 TCCAGCTTTCGCTGCGCCAGGTAAAGCAGAGACTGCCCCAAGATATTTGGGATTTGGA 5658
 Qy 2181 actgactgtagatacaaccccaaggtcaggtcaggttaagaagatgaaatagaacat 2240
 Db 5659 ACTGATTTGATATACACACCCCAAGCTCAGAGAGTACAGAGATGAATAGACCAT 5718
 Qy 2241 taagaagaccccttaacaaattgacacagagactggcaltlaatgattgagctcct 2300
 Db 5719 TAAAGAACCCCTTACCAATTTGACCACAGAGACTGCTATATGATGATGATGCTCCCT 5778
 Qy 2301 gcccttctgctctttaggtgaggaacacccctggagaagttgggtgacccctatga 2360
 Db 5779 GCCCTTTGCTTTTAAAGGTAGAGAACCCCTGACAGTTTGGCTGACCCCTATGA 5838
 Qy 2361 attgctcagagggagaccccccgttgcagaaattgctctgacatagttgatgct 2420
 Db 5839 ATTGCTTACGGGGAGACCCCCCGCTGTGTGAATTTCTCTGTACATATGCTGATGAT 5898
 Qy 2421 gctgcttccagccttcttctcagctcagcagcgctcagtggtgagcgagc 2480
 Db 5899 GCTGCTTCCAGCCTTTGTTCTTCTAGGCTCAAGCGCTGACGTGGGAGCAACGAC 5958
 Qy 2481 gtagaagaagctccggagggcctactaag---aggaacttgcaggttgcacatgct 2537
 Db 5959 GTGGAAGCAGCTCCGGAGGCTTACTCAGAGAAAGAACTTCAAGTTCCACATCGCTT 6018
 Qy 2538 ccaagcttgagattcagctatgtagagcgcacccgttgcagaaacccctcagagctggt 2597
 Db 6019 CCAAGTGSAGATTCAAGTATGTTAGAACCCAGCTGACAGAAACCTCGAGACTGGTG 6078
 Qy 2598 gaaggagactatctcgtactcttgaaccaacacagcgttgaagtcggaagatccc 2657
 Db 6079 GAAGGGCCCTTATCTGACTTTTGAACCAACAGCGCTGGAAGATGCC 6138
 Qy 2658 cactgtagatcagatcccaagcttaagccgagcgcacactcccgatctgggggtgaaagc 2717
 Db 6139 CACTGAGATCCATGCTATCCACGTTAAGCCGGCGCCACTCCGATTTGGGGTGAAGGC 6198
 Qy 2718 cgaagaagctgaaaaatcccttaagcttcgcctcagatcggttgcctactctgca 2777
 Db 6199 CAAAAAGACTGAATAATCCCTTAAGCTTCGCTCATGCGCGGTGCTTACTCTGTCA 6258
 Qy 2778 taactcctaagttaatgtaaacgcttgggaagccggaaccccaataacactat 2837
 Db 6259 TAACCTCTCAATTAATGTTAAACGCTTGTGAGAGCCCAAGCTCCATTAACCTTAT 6318
 Qy 2838 cctcaactcgttactactgactcgcgttacagatataataaagcactcaaggg 2897
 Db 6319 CTCTCACTCGTACTTACTGACTCGGTACAGATTAATTAATTAAGCACTCAAGGG 6378
 Qy 2898 aggtcccttgggagcctggtgctgaaatataatgctgcctcagatcaagtaactcctg 2957
 Db 6379 AGGCTCCCTTGGGGACCTGTGGCTGAATTAATGTGCTTGCCTGATCAAGTAATCCCTG 6438
 Qy 2958 gttcctaagtaacagucacaccccgatgactcgtgcttaaggggttaagcttgc 3017
 Db 6439 GTCTCAATACCAAGGCCACACCCCGAGATACCTCGTGTCAAGGGTTTACGTTTGCC 6498

QY 3018 caggaccaccaataatgaagaatattgtaagaatcctcaggattccttgcagcaat 3077
 |||||
 Db 6499 CAGAGCCCCCAATATGAGGAAATATGAGGAATCCTCAGAGATTCTTTTCAAGCAAT 6558
 |||||
 QY 3078 gagcgtcacaacttcaatgagtggaattggaatgagcagctctcagcaagaagag 3137
 |||||
 Db 6559 GAGAGCTGCAATACCTTAAATGATGGAATGGAATGAGGAGGAGGAGGAGGAGGAG 6618
 |||||
 QY 3138 taagtactcttctgtaaacatccacagttataatcaatattatgagcaatgga 3197
 |||||
 Db 6619 TAAGTACTCTTTTGTAAACATCCATCAAGTTATATCAATTTATATGAGGAGGAGGAG 6678
 |||||
 QY 3198 gatgagaagattggaacagcggtacaaagaatgtaagaataagaataaagcttc 3257
 |||||
 Db 6679 GATGGAAGAAGATTGGCAACAGCGGATCAAAAGATGTACGAATTAAGCAATTAAGCTGTC 6738
 |||||
 QY 3258 attcgttagacctagattacttaaaatgaattcaat 3295
 |||||
 Db 6739 ATTGCTTAGACCTAGATTACTTAAATAAGTTTCACT 6776
 |||||

RESULT 5
 LOCUS PEN133816 8918 bp DNA VRL 07-DEC-2000
 DEFINITION Porcine endogenous retrovirus type C proviral gag, pol and env genes and LTR (class B, clone 33).
 ACCESSION AJ133816
 VERSION
 KEYWORDS env gene; env protein; gag gene; gag protein; pol gene; pol protein.
 SOURCE
 ORGANISM
 porcine endogenous retrovirus.
 Viruses; Retrovirdae; Mammalian type C retroviruses; 1-Mammalian type C virus group.
 REFERENCE
 1 (bases 1 to 8918)
 Czauderna, F., Fischer, N., Boller, K., Kurth, R. and Tonjes, R.R. Establishment and characterization of molecular clones of porcine endogenous retroviruses replicating on human cells
 J. Virol. 74 (9), 4028-4038 (2000)
 2 (bases 1 to 8918)
 Tonjes, R.R.
 Direct Submission
 Submitted (04-MAY-1999) Tonjes R.R., Medical Biotechnology, Paul-Ehrlich-Institut, Paul-Ehrlich-Strasse 51-59, Hessen, D-63225 Langen, GERMANY

FEATURES
 Source
 1. 8918
 /organism="porcine endogenous retrovirus"
 /strain="type C"
 /db_xref="taxon:61673"
 /clone="33"
 /note="grown in PK15 (ATCC No. CCL-33) cell line"
 1. 707
 545
 /note="cap site"
 710..727
 /note="gly4"
 1154..2725
 /gene="gag"
 1154..2725
 /gene="gag"
 /function="retroviral capsid"
 /product="gag protein"
 join(1154..2725,2729..6313,6189..8162)
 /note="viral polypeptide"
 /codon_start=1
 /protein_id="CAB65339.1"
 /db_xref="GI:6688946"
 /db_xref="SPTREMBL:Q9Q1X5"
 /translation="MGQVTTPLSLTDHMTVEVSRANHLNVQYKKGPMQTFCASEMP
 FPDVGPSEGFNSRIILAVKAITFGTPSSHPODEYILTLWQDLADPPWPKWLN

CDS
 mat_peptide

KPKKPEPRILALGKKNKSAEKVEPSRIYELIEPPTWPEPOPPPYPAQAGVKG
 PSAPGAPVVEDEPAGTNRSGATPERDIELALILPLRTYGPMPGLOLOPYMPS
 ADLYMKTNHPFSEDEPDLTGLVESIMESHQPTDDCOLOLITLFTREERILLER
 RKNVPCADGRPTOLONEIDMSEPLTRPMDQNTNAGRESLITKRALVAGIKGASRRP
 TMLAVREVMQGNPPPSVTLERIMEARRRTPDPDTSEAKASVALALIGASLDLR
 KKLQRIEGLQIELELDLIVREKVEYVYRRETEEEKQOREKEREREERDRDRKNTL
 KILAAVGEKSSREREDEFRIRSGPPOSGMGNMTPLDKQOCAYCEKGMANCK
 KNGKPKVTLLEEKDGRSGSDPLPEPVTTLKVGQPEVPTLVDGASLPLDPLGLT
 KEKSMWAGATGOORYPWTRRTVDLAVGRTHSELVIPCPOVATLITLQGLT
 SPEQGRPEVSNNKPIYTLITOLDDEYRLSPQYKQPDQIOSMIEOPQMAETKAGG
 LAKQPPVOYLKASATPVSVKQPLSREAREGILPVOHRLIQGLILPVQSPMNPPL
 LPKRVGINDRFPVODLEKVNKRVODLHPVVPNPNLNLALPEPMYVTLQDKAF
 CLRLHPTSPLEFEMRDGPGRGTGLTWRLPLPOGFENSPITIDEALRLANERIOH
 POVTLIQYVDDLLAGATKODLEGTALLLLEISDLGRASAKKAOICREYVLYGS
 LSGORMLTEARKKTVVOIPAPTRKQVREITGAGTCRMTIGEPATLAPPLPKRE
 KGEPMAPENHOKAPDAIKKALLSPALADLVYKFPPTLYDEKRGAKGLVTLTLPW
 RPPVAVLYSKLIDPVASGMPVCLKALAAVALLVDADKLITLIGONITVYAPHALENIVRO
 PDRMTNARKTHYQSILTERVTEFAPPALNPNATLLPERTEDEPVHDCQLIEETG
 VKKDLITDILPTEGVLTWETDSSYVEGKRMAGAAVVDGTHITWASSLPGTSQAKAE
 LMALTOALRLAEKKSINITYDSRYAFATAVHGAITYKORGLITSAGRETKNKEITSL
 LEALHPKRLALIHCPGHOKAKDLISRGNDMPDVAQAQANLPIITETPAPER
 ROYILEDMQELKTIIDQSETPETCTYTSYKEILPHKEGLETVQULHRLITETKHLQ
 OLVRTSPETHVRLPGVADSVVHCPCQILVNANPSRLIPKRLRGSHPGAMHVEDTE
 VKPAKYGKYLIVLVEVDFSGVWEAVPRKLETSTVYAKKILIEETFPFGILPKVGSNG
 PAFPAVQSGSLAKTILGIDMKLHCAYRPOSSGOVEPMRRTIKETLTKLTETGINDIA
 LLPLVLEFRVNTPGOFGLTPYELLGGPPPLVEATSVHSDVLLSOPLSRKALBMY
 KORAMKOLREAVSGEGDLOVPHROVDVSVYVRHAGNETETWKRPVLYLTTPPVA
 KVEGISTWIIHSHVKIAPPPDSGGRAKTENPILRLHLRLVPIYNNNSPQULPTLSM
 RHLPTRGGERKRLPIPLFASIAFELTLTTPPAASSKRLIDSSNPHRLPLTWLITDP
 DTGVTAVSTRGVARPTWPELHPCLELIPAVKSTPNPNRISYGFCCPGTEKERYC
 GDSESEFCRMSVTSNDGMKWPISLQDVKSKFPVNSGKVKVKKLKDCKSCPSD
 LDYLKISTFEKQENIOWKINGSMWIVYKKGSGSTLIRLITPENGTEPPVAG
 PDKVLAEQPPALEPPEPNILPOLYTSIRPITPOENSGTGLIPTPTPNSSQVPAKT
 GORLESLIOGAFQAINSTDPTDSSCLSSPPRYEGBAKGKNVTRKHNQCTW
 GSRKILITLESKGTCTGAPPSHOLCTIVYVYQDASNNQILVGYRMAACNGL
 TPCVSTVFNOSKPCVWQIVPRVYHPEVYVDEVDYASNNQILVGYRMAACNGL
 LGTVAGVGTGALALITGPOOLEKGLGFLVHAMIEDIALAESVNLSESLTGSYVL
 QNRGADLPLELREGGLCALKECCFCEVHSGAIDRMSKLRRLERLRERREEDQW
 FEGWENSPMNTTLLSLATGPLVLLLLLVGSCILINRVAIVREVSVAQIMVLRQO
 YQGLISGSETDL"
 2729..6313
 /gene="pol"
 /function="retroviral protease and polymerase"
 /product="Pol protein"
 2729..6313
 /gene="pol"
 6189..8162
 /gene="env"
 /function="retroviral envelope"
 /product="Env protein"
 6189..8162
 /gene="env"
 8212..8918
 /note="env"
 8811..8816
 LTR
 polyA_signal
 BASE COUNT 2327 a 2245 c 2334 g 2012 t
 ORIGIN

Query Match 53.3%; Score 1769; DB 14; Length 8918;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 2219; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 64 cttttgcttgaatgagagatccaggtacgggagaagcggagcctcagtcgacc 123
 |||||
 Db 3581 CTTTITTCCTTCAGAAATGGAAGATCCAGGTACGGGAAGACGGGAGCTGACAC 3640
 |||||
 QY 124 cgaatgcccccaagaggttcaagaactcccgacacatttgagcaagagcctcaagggagc 183
 |||||
 Db 3641 CGAGTGGCCCCAAGAGGTTCAGAAAGCTCCCGACCACTTTTGACGAAGCCCTTACACAGGGAGC 3700
 |||||
 QY 184 ctggccaacttcaagatccacaacccctcagtgagcctcctcagtaagtgatgacctg 243
 |||||
 Db 3701 CTGGCCAACCTTCAGAGATCCACACCCCTCAGTGACCCCTTCACATACGTGATGACCTG 3760
 |||||

QY 244 ctctgctgagagccacccaacagactgcttaagaagtaacgaagcaactgctgtaa 303
 |||||
 Db 3761 cttctggcgaggccaccacacagactgcttaagaagtaacgaagcaactgctgtaa 3820
 |||||
 QY 304 ttgtctgacccaggtctaaagagctctgtaagaagcccaagatttgcagaagaagta 363
 |||||
 Db 3821 ttgtctgacccaggtctaaagagctctgtaagaagcccaagatttgcagaagaagta 3880
 |||||
 QY 364 acatacttggggttaacagtttgcgggagcgagcatgctgtaagcgagcgagcaaa 423
 |||||
 Db 3881 acatacttggggttaacagtttgcgggagcgagcatgctgtaagcgagcgagcaaa 3940
 |||||
 QY 424 actgtaagtcagataccgagccccaacacagcaaaagatgagaagtttttggggaca 483
 |||||
 Db 3941 actgtaagtcagataccgagccccaacacagcaaaagatgagaagtttttggggaca 4000
 |||||
 QY 484 gctggaatttgcagactgtaaccggggttgcagacttaagcagcccaacttaaccg 543
 |||||
 Db 4001 gctggaatttgcagactgtaaccggggttgcagacttaagcagcccaacttaaccg 4060
 |||||
 QY 544 ctacccaagaagaaggggagatctcctggtctcctgagcaacgaaggaatttgaatgct 603
 |||||
 Db 4061 ctacccaagaagaaggggagatctcctggtctcctgagcaacgaaggaatttgaatgct 4120
 |||||
 QY 604 atcaaaaagggccctgtaagcgacactgctctggtccctcctgagcaataaaccctt 663
 |||||
 Db 4121 atcaaaaagggccctgtaagcgacactgctctggtccctcctgagcaataaaccctt 4180
 |||||
 QY 664 acccttaagtgatgagagtaagaaggaagcccgagagatttaacccaacccataga 723
 |||||
 Db 4181 acccttaagtgatgagagtaagaaggaagcccgagagatttaacccaacccataga 4240
 |||||
 QY 724 ccatggaagagagcctggtgctcactgctcaagaagagcttgctcgtgaagcagtggtg 783
 |||||
 Db 4241 ccatggaagagagcctggtgctcactgctcaagaagagcttgctcgtgaagcagtggtg 4300
 |||||
 QY 784 cccgtatgctgaaggactatcgaaactggtgacatactggtgaagagcgtgcaaaatg 843
 |||||
 Db 4301 cccgtatgctgaaggactatcgaaactggtgacatactggtgaagagcgtgcaaaatg 4360
 |||||
 QY 844 acttgggagagataatactgtaataagcccccatgattggagaataatgctggagag 903
 |||||
 Db 4361 acttgggagagataatactgtaataagcccccatgattggagaataatgctggagag 4420
 |||||
 QY 904 ccccgagacagatgataagcaaacgcccagcatgacccaatacaaacgctgcttcaaca 963
 |||||
 Db 4421 ccccgagacagatgataagcaaacgcccagcatgacccaatacaaacgctgcttcaaca 4480
 |||||
 QY 964 gaagaggtaacttgcctcaacagcgtctcaaccctgcaactctctgctgtaagag 1023
 |||||
 Db 4481 gaagaggtaacttgcctcaacagcgtctcaaccctgcaactctctgctgtaagag 4540
 |||||
 QY 1024 acgataagacagatgataagcaaacgcccagcatgacccaatacaaacgctgcttcaaca 1083
 |||||
 Db 4541 acgataagacagatgataagcaaacgcccagcatgacccaatacaaacgctgcttcaaca 4600
 |||||
 QY 1084 aaggaacttaacagataccgctgactggaagaatgtaactggttcaactgtaaggaagc 1143
 |||||
 Db 4601 aaggaacttaacagataccgctgactggaagaatgtaactggttcaactgtaaggaagc 4660
 |||||
 QY 1144 agctatggtgtaagagtaagatggtctggggagcgagtggttgaacggagcccgacg 1203
 |||||
 Db 4661 agctatggtgtaagagtaagatggtctggggagcgagtggttgaacggagcccgacg 4720
 |||||
 QY 1204 atctgggcaagcagcctgaggaaggaacttgcagcaaaagcttgaagtaatggccctc 1263
 |||||
 Db 4721 atctgggcaagcagcctgaggaaggaacttgcagcaaaagcttgaagtaatggccctc 4780
 |||||
 QY 1264 acgcaagccttgcggtgcgcgaaggaataaccataacatttatacggacagcggtat 1323
 |||||
 Db 4781 acgcaagccttgcggtgcgcgaaggaataaccataacatttatacggacagcggtat 4840
 |||||

QY 1324 gcccttgcagctcacacgtaacacggtccatctataaaagaaggggttgccttaacca 1383
 |||||
 Db 4841 gcccttgcagctcacacgtaacacggtccatctataaaagaaggggttgccttaacca 4900
 |||||
 QY 1384 gcaaggaggaataaagaacaaagaaattcctaagcctatataaagccttaacttg 1443
 |||||
 Db 4901 gcaaggaggaataaagaacaaagaaattcctaagcctatataaagccttaacttg 4960
 |||||
 QY 1444 ccaaaaagctagctatataacactgctccgtgagacaataagaagccaaagatcatalct 1503
 |||||
 Db 4961 ccaaaaagctagctatataacactgctccgtgagacaataagaagccaaagatcatalct 5020
 |||||
 QY 1504 agaaggaaacagatgctgtaacgggttgcgaagcagcagccagcgtgttaaccttgc 1563
 |||||
 Db 5021 agaaggaaacagatgctgtaacgggttgcgaagcagcagccagcgtgttaaccttgc 5080
 |||||
 QY 1564 cctaataaagaagcccaagcccaagcccaagcccaagcccaagcccaagcccaagccca 1623
 |||||
 Db 5081 cctaataaagaagcccaagcccaagcccaagcccaagcccaagcccaagcccaagccca 5140
 |||||
 QY 1624 caagagataaaagataagacagatctctgagactccggagggagggacttaacctca 1683
 |||||
 Db 5141 caagagataaaagataagacagatctctgagactccggagggagggacttaacctca 5200
 |||||
 QY 1684 tatgggaaggaatctcgtcccaagaagaaggttgaataatgtccaaacataatcgt 1743
 |||||
 Db 5201 tatgggaaggaatctcgtcccaagaagaaggttgaataatgtccaaacataatcgt 5260
 |||||
 QY 1744 ctacccaactagagaaactaaacacactgagcagttgtgcagaatcccttaactatglt 1803
 |||||
 Db 5261 ctacccaactagagaaactaaacacactgagcagttgtgcagaatcccttaactatglt 5320
 |||||
 QY 1804 ctgaagactacaggaagtggtgactcgtggttcaaaactgtgtgcccctgcagctggt 1863
 |||||
 Db 5321 ctgaagactacaggaagtggtgactcgtggttcaaaactgtgtgcccctgcagctggt 5380
 |||||
 QY 1864 aatgctaaactcctccagaataactcctcagaagaagactaaagggaagcccaacggcgt 1923
 |||||
 Db 5381 aatgctaaactcctccagaataactcctcagaagaagactaaagggaagcccaacggcgt 5440
 |||||
 QY 1924 cactgggaagtgacttaactgagtaagcccgctaaatacgcgaataataatctatgt 1983
 |||||
 Db 5441 cactgggaagtgacttaactgagtaagcccgctaaatacgcgaataataatctatgt 5500
 |||||
 QY 1984 gttttgtgaacacttcaagatggttgaaggttaactcactaaagaagaactca 2043
 |||||
 Db 5501 gttttgtgaacacttcaagatggttgaaggttaactcactaaagaagaactca 5560
 |||||
 QY 2044 accgtgtgctaaagaataactggaaggaatlttccaagaatttgaataactaaagta 2103
 |||||
 Db 5561 accgtgtgctaaagaataactggaaggaatlttccaagaatttgaataactaaagta 5620
 |||||
 QY 2104 atagggtcagacaatggttccagcttctgtgcccggttaagtcaagagccttgcgaagta 2163
 |||||
 Db 5621 atagggtcagacaatggttccagcttctgtgcccggttaagtcaagagccttgcgaagta 5680
 |||||
 QY 2164 ttggagatgataagaaactgcatgctgatacaaaccccaagtcaggaacagtagag 2223
 |||||
 Db 5681 ttggagatgataagaaactgcatgctgatacaaaccccaagtcaggaacagtagag 5740
 |||||
 QY 2224 aggatgaatagaacccaataagaagcccttaacaaattgacccaagagaactggaataat 2283
 |||||
 Db 5741 aggatgaatagaacccaataagaagcccttaacaaattgacccaagagaactggaataat 5800
 |||||
 QY 2284 gattgagt 2291
 |||||
 Db 5801 gattgagt 5808
 |||||

RESULT 6
 PER17013 PER17013 7808 bp DNA VRL
 LOCUS Porcine endogenous retrovirus gag, pol, env genes. 14-APR-2000
 DEFINITION

ACCESSION Y17013
VERSION Y17013.1 GI:3005985
KEYWORDS env gene; gag gene; pol gene.
SOURCE porcine endogenous retrovirus.
ORGANISM Porcine endogenous retrovirus.
Virus; Retrovirus; Retroviridae; Mammalian type C
retroviruses; 1 mammalian type C virus group.
REFERENCE
AUTHORS Czauderna, F., Fischer, N., Boller, K., Kurth, R., and Tonjes, R.R.
TITLE Establishment and characterization of molecular clones of porcine
J. Virol. 74 (9), 4028-4038 (2000)
JOURNAL
MEDLINE 20219394
REFERENCE 2 (bases 1 to 7808)
AUTHORS Toenjes, R.R.
TITLE Direct Submission
JOURNAL Submitted (30-MAR-1998) R.R. Toenjes, Paul-Ehrlich Institut,
Paul-Ehrlich Str. 51-59, P.O.Box 1740, D-63207 Langen, FRG
FEATURES
source
1. 7808
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
/tissue_type="kidney"
/cell_line="PK15"
/gene="gag"
260..1834
/gene="gag"
260..1834
/codon_start=1
/protein_id="CAAF6583.1"
/db_xref="GI:3005986"
/db_xref="SPTREMBL:O73504"
/translation="MGQVTPPLSLTLHDMWTVSRANLISVQKKSGWQPCSEMP
KPKPPRILALGKRRKHAKEVSPRIPEPEPPPPVQPPVQAGAVR
PSAPGAPVGEAGAGTSGRATGATDELALPLRTGPPMPGQLOPQVFFS
ADLYNKTNHPPESEDPRQLGLVESLMEHSHOPTDDCOQLQTLTAEERILL
RKNVGPADRPOLONEIDMGPPLRLGMDVYTAGRSLSIKYRQALVAGKGRSP
TNLAKREVWQGNPEPSEVLEIRMEARPRPTPTPTSEAKASALAFITQSLDTR
KLORLEIGQAEIPLDLYREAKATYGTETBEKREKEEREERDRORKNLT
KILAAVEKSSREBRDRKTRSGPRSGNLGNTPLDKQACVCKKGMANCPK
KGNKGPVLALEDKD"
2120..4726
/gene="pol"
2120..4726
/gene="pol"
/codon_start=1
/protein_id="CAAF6582.1"
/db_xref="GI:3005987"
/db_xref="SPTREMBL:O73505"
/translation="MGAQISFEGCPPEVSNKPTIVTLTLDDEYRKYSPQKPPD
TOSMLEPQQAETAGMGLAKQVPOVITLAKSTVSVQYPLSFRAGTWPVQ
RLIQGLIVPQSVNTPPLLEPRKPTNDYRPOVLDVSNKRVODIHPVFNPLLS
ALPERNMVTVLIDKDAFCLRLHETSQLPFAFMRDPTGRTGLTWTRLPQCFKS
PTTDEALHRLDRNRIQHPQVTLTQYVDDLLAGATKODLETKALLLESLGR
ASAKAOICRREVLYGLRGGORWLTAEARKVVOIPAPTTAKOYREFLTAGGER
LWIPGATVAILPLTKEGSEFMAHOKAPAIKALISAFALAPVTRFTLY
VDERGVARCVLTQTLGPRRPVATLSKKLDVASGVCALIAVAALIVKADKLT
LQONTVIAFPALENTVROPEDRMNTNARMTHYOSLLTERVTFAPPAALNPATLLE
ETDEPVTDCQOLLEETGVKRLDLDLPLTGEVLTFTDSSVYVEGRBAGAAVVDG
TPTIWAASSLPETGSOKAEMALTOALFLAEGKINITYDSRYAFATVHGAIVYOR
GLTASAGREIKKKEEILSLLEALHLPKRLATIHGPGHOKADLISRGONADVAQA
AOAVNLPLIETPAPEPRROYTLEDQETIKIDQSETEBGTCTYTGELPHKEG
LKVVOIHRITHTICKHQQLVRTSPVHVLPLPVADSVAKHCPQOLVANSRLPP
GKRLAGSHPGAHWEVDTEVPKATYGNKYLIVFDTEGWEAVPTKKTSTIVAKY
WKRFRLLEYLR"

gene
CDS

5294..7267
/gene="env"
5294..7267
/gene="env"
/codon_start=1
/protein_id="CAAF6583.1"
/db_xref="GI:3005988"

/db_xref="SPTREMBL:O73506"
/translation="MHPILSWRLPMPGPKRLRIPLSPASTAMFLTITLTPQASSK
RLIDSSNPHRLSLTILIIDPDGVYVNSTRVAPRGMTWPELHFLRLINPAVKSTP
PNLVRSTGYTCRGTETKERTGGSESEFSCRWCVISNDQMKPILSDRVKFEVN
SGPKRYVMKLYKDKSCSPSDLDLYLAKISFEKGRONIKRMINGMSMGIVFYKGGGA
GSTLPIRLIETGTEPEPVAAGPKVLAEGPPALPEPHNLVPOLSLPLDITPOPSN
GTGLIPTNTPRNSPCVPTGTGRLSTLISOGAOLINSDPATSSCNILCSGPPRY
EGMAKEKENVTKENHROCTWGSRNKLTILESCATCTGKAPSHQILCYSTVVEO
ASENQTLVPGTNMMACTNGLTPCVSSSVFNOSKDECVWQIVPVRVYHPEVDEY
DYRNKRPKREPVSLTILAVMIGLTAAGVGTAPALITGPOLEKIGELHAAFTEDLR
ALEESVSNLEESLSLEVIQNRGLDILFLREGGLCAALKECCFVYDHSALRDS
MSKLEERLERREREADQGEFGEWFGNPSRPMWTTLSALTGLVLLILLTVGPCLIN
RFVAFVERGAVOIMVILNROVGLISQETDL"
BASE COUNT 2077 a 1971 c 2047 g 1713 t
ORIGIN
Query Match 50.6%; Score 1680; DB 14; Length 7808;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2030; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 21 agatgctctctctgctgagattacaccccaagcatttggccttgatg 80
Db 2644 AGATGCTCTCTTCTGCTGAGATTACACCCACTAGCCACACTTTTGGCTTGAATG 2703
QY 81 gagaatccaggtacggaagaacccggaagtcacccgagaccgagcccaaggtt 140
Db 2704 GAGGATCCAGGTACGGGAAGACCGGGACGTACCTGGACCCGACCCCAAGGGTT 2763
QY 141 caagaatcccccagcaatttgaagaagccctacagaagccctggccttaagat 200
Db 2764 CAAGAATCCCCGACCACTTTTGAAGAGCCCTACAGGACCTGGCCCACTTGAAGAT 2823
QY 201 ccaaacctcagtgagtcacccctccagtcagtcagtcacccctctgaggaagccac 260
Db 2824 CCAACACCTCAGTGAGTACCTCTCCAGTACGTGATGACCTCTTGGCGGAGCCAC 2883
QY 261 caaacgagctgcttagaaggtacgaagcactactgctgaatttgcacttagtca 320
Db 2884 CAACAGAGCTGCTTAGAAGGTACGAAGGACACTGCTGGAATTTGACCTTAGGCTA 2943
QY 321 cagaagctctgcttagaagaagcccaagatttgcaggaagaaggttaactattgggtacag 380
Db 2944 CAGAGCTCTGCTAAGAGGCCCCAGATTGCAAGAGAGGTAACTACTTGGGTACAG 3003
QY 381 ttgcagggcggcagcagatgctgacggaggaagaagaactgtagttcaatcc 440
Db 3004 TTGCGGGGGGGGACGAGGCTGACGAGGACGAGGAGAAACTGTAGTCCAGATTAC 3063
QY 441 gcccacacacacacgaagaagaagtttgcagagctggaatttgcagact 500
Db 3064 GGGCCCAACACACGCAACCAAGTGAAGAGTTTGGGAGACGCTGGATTGGAGACT 3123
QY 501 gttgattccgggttggagaccttagcagcccaacttaccgcttaacaaagaagaag 560
Db 3124 GTGGAATCCGGGTTTGGCACTTAGCACCCCTTACCCGCTTAACCAAGAAAGAG 3183
QY 561 ggaattctcctgggctcctgagcagcagaagaagcatttgaattacaaagaagccctct 620
Db 3184 GGAATTCCTCTGGGCTCCGAGACACAGAGGACATTGATGATCAAAAAGGCGCTGCT 3243
QY 621 gaggcagctgctctggcctccctgaagtaactaaaccccttaccctttagtgaatga 680
Db 3244 GAGCGACCTGCTGCTGGCTCCCTGACGACTTAACCCCTTATGAGGATGA 3303
QY 681 ggcgaaggaagtagcccgaggagtttaacccaaccccttaggacatggaagaagcctgt 740
Db 3304 GCGTAAAGGAGTAGCCCGAGGAGTTTAAACCAACCTTAGGACATGAGAGAACCTGT 3363
QY 741 tgcctactctcaagaagcctgactcctgtaacagtggttggcccgatgctgaagggc 800
Db 3364 TGCTACTCTGCTCAAGAACCTTGATCTGTAGCAGTGGTGGCCGATATGCTTGAAGCC 3423

QY 801 tatcgagctgtgcccatactgttcaaggaagcgttacaattgactttggagacagatat 860
 |||
 Db 3424 TATCGAGCTGTGGCTATCTGTCAAGAGACGCTACAAATTTGACTTGGAGACGAATAT 3483
 QY 861 aactgttaaatgcccccatgcatgttgagaacactgttcgagagcccccaagccatgcat 920
 |||
 Db 3484 AACTGTTAATATGCCCCCATGATGAGAACTCTTGGCAGAGCCCCCAGACCATGAT 3543
 QY 921 gaccagagccgcaatgagaccactatcaaaagctgttctcaagagagaggttcaattcgc 980
 |||
 Db 3544 GACCAACGCGCCGATGACCCACTATCAAGCCCTGCTTCTCAGAGAGAGGCTTCTGCG 3603
 QY 981 tccaccagcgcgtctcaaccctgcacactcttctgcttcaagaagactgtgaacagctgac 1040
 |||
 Db 3604 TCCACACAGCCGCTCTACACCTTCGCACTTCTTGGCTGAAGAGATGATGAACAGTGAC 3663
 QY 1041 tcatgtttgcatcaactttgatttgaagagactgttggttcgcaagaaaccttcaagacat 1100
 |||
 Db 3664 TCATGATTTGCACTCACTATTGATTTGAGAGAGACTGGGGTCCGCAAGGACCTTACAGACAT 3723
 QY 1101 accgctgactgtgagagactgttaacctgttcaactgttcaagagagactgttgtgtgagag 1160
 |||
 Db 3724 ACCGCTGACTGGAGAGTCTTAACCTGTTCTCTAGAGAGACACTTATGTTGTGAGAGG 3783
 QY 1161 taagagagatggtctgagagcgagtggttgaagcggaaccgcaagactgtgagcagagct 1220
 |||
 Db 3784 TAAGAGAGATGCTGGGGCGGCGGTGTGTGAGCGGAGCCACAGATCTGGCGCAGACGCT 3843
 QY 1221 gccggaagaaacttcaagcgcaaaagctgagctcaatgagcctcaagcaagctttggcgt 1280
 |||
 Db 3844 GCCGGAAGGAATCTTCAAGCCCAAAAGGCTAGCTCAAGGCGCCACAGCAAGCTTTCGCGCT 3903
 QY 1281 gccggaagaaacttcaataacattatatacgagcagagctgttgccttgcagactgca 1340
 |||
 Db 3904 GCCGGAAGGAATCTTCAACCTTTATATGAGAGAGAGAGAGTGTGCTTGGCAGCAGCA 3963
 QY 1341 cgtacaggggcatctataaacaaggggttgccttaccctcagcagagggagaaataa 1400
 |||
 Db 3964 CGTACAGGGGCGCATCTATTAACAAGGGGTTGCTTACCTCAGAGAGGAGGAAATAAA 4023
 QY 1401 gaacaaagaaagaaattctagccttatagaagccttgaacttgcacaaagagctagcat 1460
 |||
 Db 4024 GAACAAAGAGAAATTTCTAAGCCTATTAGAGCCTTACATTTGCCCCAAAAGGCTTAGCTAT 4083
 QY 1461 tatatactgctctgagacatcaagaagcacaagatcatalatcagagagaaacagatgac 1520
 |||
 Db 4084 TATACACTGTCCTGAGACATCAGAAACCAAGATCATATCTAGAGGAGAACAGATGAC 4143
 QY 1521 tgacgggttgcgaagcagcagcagcagctgttacccttgcctataatagaagcgc 1580
 |||
 Db 4144 TGACCGGCTTGCACAGAGGACGCCAGCTGTTAACCTTCTGCTTATTAATGAAACGCC 4203
 QY 1581 caaagcccaagaaaccagagcagcagcagcagcagcagcagcagcagcagcagcagc 1640
 |||
 Db 4204 CAGAGCCCCAGAAACCCAGACGACATACACCTAGAGAGCTGGCAAGAGATTAAGAAAT 4263
 QY 1641 agacagcttctctgagactcaggaagggagcgtgctatacctcatatbggaagaaatcct 1700
 |||
 Db 4264 AGACAGCTTCTGAGACTCCGAGAGGAGACTCTTATACCTCATATGAGGAAGAAATCCT 4323
 QY 1701 gcccccaaaagaggttataatattccaagagataacttcttaaccactcagagac 1760
 |||
 Db 4324 GCCCCAAAGAAAGGGTTAATAATATCTCCAAACGATATCTCTCCACCTAGGAAC 4383
 QY 1761 taacacactcagcagacttgcagaaacatcccccttcaatgttctgaagcctcagagact 1820
 |||
 Db 4384 TAAACACCTCGACAGATTGCTCAGAACATCCCTTATCTATGTTCTGAGGCTCCAGGACT 4443
 QY 1821 ggtcgtacgtggtgcaaaatattgttgccttgcagcagctgtttaaagtcaatcctcag 1880
 |||
 Db 4444 GGCTGACTCGTGTGATTAACATATGTTGCTGCCCTCCAGCTGTTAATGTAATCTTCCAG 4503
 QY 1881 aatacctccagaaagagactaagggaagcaccagcgctcactcgtggaagtgtgactt 1940

Db 4504 AATACCTCCAGAAAGAGACTAAGGGAAGCCACACAGGCGCTCAGTGGGAAGTGACTT 4563
 |||
 QY 1941 cactgaggttaagcccgcttaataataggaacaaatattatgttttttgaagcactt 2000
 |||
 Db 4564 CACTGAGGTAAAGCCGCTTAATTAAGAAACAAATATCTATGTTGTTGTAAGACACTT 4623
 QY 2001 ttcagatgtgttagagcttactcactaagaagagactcaacgctggtgtgctaa 2057
 |||
 Db 4624 TTCAGATGCTGTAAGGCTTATCTTACATAGAAAGAGACTTCAACCGGTGGCTTAA 4680

RESULT 7
 PEN293657
 LOCUS
 DEFINITION
 Porcine endogenous type C retrovirus proviral gag gene, pol gene and env gene, class B, clone 213.
 ACCESSION
 AJ293657
 VERSION
 1
 KEYWORDS
 env gene; envelope; gag gene; group specific antigen; pol gene; polymerase; polyprotein.
 SOURCE
 ORGANISM
 Viruses; Retroviral viruses; Retroviridae; Gammaretrovirus.
 REFERENCE
 1 (bases 1 to 8763)
 Krach, U., Fischer, N., Gzauder, F., and Tonjes, R.R.
 Comparison of replication-competent molecular clones of porcine endogenous retrovirus class a and class b derived from pig and human cells
 JOURNAL
 MEDLINE
 2 (bases 1 to 8763)
 REFERENCE
 2 (bases 1 to 8763)
 Tonjes, R.R.
 Direct Submission
 Submitted (31-JUL-2000) Toenjes R.R., Medical Biotechnology, Paul-Ehrlich-Institut, Paul-Ehrlich-Str. 51-95, Langen, 63225, GERMANY

FEATURES
 source
 location/Qualifiers
 1..8763
 /organism="porcine endogenous type C retrovirus"
 /proviral
 /specific_host="Sus scrofa"
 /db_xref="taxon:70540"
 /clone="213"
 /cell_line="PK15"
 /note="class B"
 1..630
 467
 /note="cap site"
 633..650
 1077..2651
 /gene="gag"
 1077..2651
 /gene="gag"
 /codon_start=1
 /product="group specific antigen"
 /protein_id="CAC38625.1"
 /translation="MGQVTTPSLTLDHMTVEVSRANHLISVOKKGPWOTFCASMP
 PEDVMPSEGFENSEILAVKAIIFGSGSPDEEPIILWDLAEPPVWPKPMLN
 KRPKGPRIHLEGNKHSKAKVPSGTYEIEPEPTWEPDPPVPAAGAVAG
 PSAPPGAVYVEBPAGTSSRGATPERPDELAIIPLTYGPPMGGLDLPWPSS
 ADLVWKNTHPEPSEDPPRLTGLVESLHFSHPWPDDCOQLLTFTTEREKLIEA
 RNKVPDAGRPOLONEIDMGFPPLTRPGWDNTAEGRESUKIYKQALVAGLRASRR
 TMLAVREVMQPNPSPVFLERLMEAFRRTPDPISQAOKASVALATIGQSLDIR
 KTLQREGLOEALDIYRAEKVYVRETDEEKREKEREEREREERDRROEKNTL
 KTLAAVVEKSSRERERDRFRKIRSPRQSGNLGRPTLDKQCAVCKEKHMANCPK
 KENKPKVLAIEDKD"
 2652..6239
 /gene="pol"
 <2652..6239
 /gene="pol"
 /note="complete pol ORF, gag-pol is expressed as a

		polypeptide	
		/citation=[1]	
		/codon_start=1	
		/product="polymerase"	
		/db_xref="GI:14275843"	
		/protein_id="CAC39626.1"	
		/translation="GRGSDPLPEPRVTLKVEGQVPEFLVDGAEHSLDLOPLKKE	
		KSMWATGQROYPWTRRTYDLAVGVTSHSVIPECPVLLGRLLTKMGAQISF	
		KQVPEVSNKPIVTLTOLDDELRVSPQVRCVDDOISMLRQPMANETAGMGLA	
		KQVPOVIOKASATPVSVROPLSRREGIMVHROILLQOGLTIVVQSPKPTPLP	
		VRKSTNDYRPVODLRVKNRYVODIHPVNPVYLISALPERKMTVVLVDSKPFOL	
		RLHPTSOPLPEPRMDPEGTGDTWRLPQEGKNSPEFDEALHDLNFIQHQ	
		VTLQYVDLLAGATKODLEGTALLLEISDYGASAKAKQICRETVYIGSLR	
		GGQRLTEARKKATVVOIPAPTAQVREELGTAFCLMIPGKRLRGSHGAEHNPTEV	
		ESMAPEHQAQFADIKALLSPALAPALVDEPTLVDKRGVAVRVLNQTGKRR	
		PVAYLSKIDPVAAGMPVCLKAIAVAALIVKADKILGONIVYLAHLENTVROP	
		DRMTNARHTYOSILLTERVTPA PPAIINPAILPEPTDEPTVHCHOLLIEEVR	
		KDIDIPITGVLVITFDGSSVYVKKMGAAYVDGTRITMASLPEGTSAKAEIM	
		ALTOALRLAEKSTINITYDSYAFATNAHVQATYKQGLITSGARETKNEELISLE	
		ALHLEKRLAIIHCHQKQKAKDLISRGQMDRVAKQAQAVNLPILEPKAPEPRRO	
		VTLDEMDKIKIDQFSETPGCTYSYKELIPKELLEVVOGIRHLTGTHLOOL	
		VTSFVYHVLRLPVADSVYKHCPCOLVANNPSRIIPGKRLRGSHGAEHNPTEV	
		PAKYGNKILVEVDTPFGWVAVYPTKETSIVYAKKILEIIFREGIPVIGSDNPA	
		FVAQVSGILKILGIDMKLHCAYRPOSSGVERNRRTIKETLTKLTETGINDMAL	
		PVLPRRANTPGQGLPYELTLEGGPPVLEIASVHSADVLISQPLSRKALEVRQ	
		RAMQOLREAVSHGKIDQVPHRFQVGDVYVRHAGLMELETRMGKPYLVLTPTAVK	
		EGISTWTHASHVGLAPPDSGWRKAEKTENPLKRLHRLVYVSNNSPGQ"	
		6112. .8085	
gene		/gene="env"	
		6112. .8085	
CDS		/gene="env"	
		/codon_start=1	
		/product="envelope"	
		/protein_id="CAC39627.1"	
		/db_xref="GI:14275844"	
		/translation="MPTLSMRHLPTRGSGPKRLRIPSPASIAMLITLTTPQSSK	
		RLIDSSMHPRLSTLTLIDPDTGVYVNSIRGVAPGTWPELHCLRLINAVSTP	
		PNIIVRSYGFCCCTEKEKTCGGSGFSFCRMSCVTSIDGRKPFSEVN	
		SGRQKIVMKLYDKSCSPSDLYLKISPEKQKQNIQKWINGMSQIVFVKYGGGA	
		GSTTLIRLETETEPVAPGPKVLAEQPALEPHNI.PVQLSLPESLDTGPPSY	
		STGLIPLTPNPSGPVYKQRLFSLIQGAQALINSDPATSSCWLCSGGPPY	
		EGMAKEKFWVTEKHNOCQWTSRNLTLTEVSGKQCTGKAPSHOHCNSTVYEO	
		ASENOYLVGVNMMACNTGIPPCVSTFVFNOSKDCVVOIVPRYVIEPEVILDEY	
		DYRYNRKRPVSLTLIAMLGLSTAVGTGIDALITGQLEKLSIGELHMAVTDLR	
		ALBESYNLEESLTLSLEVVLNRGLDILFLEGGICALEKECCFVYDHSGARIDS	
		MSKILRELERRRERADQGEFGMNRSPMTTILSALTGLVLLLTITVPCLLIN	
		RFVAFERERYSAVOIMVLRQYGLISQGETL"	
		8133. .8763	
LTR		8133. .8763	
BASE COUNT		2301 a 2200 c 2287 g 1974 t 1 others	
ORIGIN		8655. .8660	
Query Match		50.0%; Score 1659; DB 14; Length 8763;	
Best Local Similarity		99.5%; Pred. No. 0;	
Matches 2259; Conservative		0; Mismatches 12; Indels 0; Gaps 0;	
QY	21	agatgcttcttctgctgataaaccacttaagcaacacttctgcttgaatg	80
Db	3461	AGATGCTTCTTCTGCTGAGATTACACCCCTAGCCAAACCTTTTGGCTTCGAATG	3520
QY	81	gaagatccagatcaggaagaccggagcaccctggaacccgacttgccttgaatg	140
Db	3521	GAGAGATCCAGATCAGGAGAACCGGCGACTCTGACCTGAGCCCGCAAGGGGT	3580
QY	141	caagaactcccgacacacttctgaagcaagcctacacaggaacttgccttgaatg	200
Db	3581	CAGAAGCTCCCGACACTTTCTTGACGAAGCCCTACACAGGAGACTGGCCAATTGAGAT	3640
QY	201	ccaacacccctcaggtacccctcctcagtaagtgatgactccttctgaggaagcaac	260
Db	3641	CCAACACCTCTAGGTACCTCTCCAGTAGTGATGACCTGCTTCTGGCGGAGCCAC	3700
QY	261	caaacagagctgcttgaaggtacgaagcactactgcttgaatgcttgcacttaggca	320
Db	3701	CAAAACAGACTGCTTGAAGAGTACGAAGGACACTGCTGGAATGCTGACCTAGCGTA	3760
QY	321	caagaccccttgcctgaagaagcccaagatttgcaggaagaggtgaacatacttgggtacag	380
Db	3761	CAGACCTCTGCTTAAGAAAGGCCAGATTGCGAGGAGAGATTAACATCTGGGGTACAG	3820
QY	381	tttgcagggcgagcgacgatactgctgacgagacaggaagaacttgaatccataacc	440
Db	3821	TTTGCGGGGCGGGCAGCGATGCGTACGAGGACAGGAAGAAACCTGTATGTCAGATACC	3880
QY	441	ggccccaacacacagcccaacagttgaagagtttttggggagacgttgaatttgcagact	500
Db	3881	GGCCCAACACACAGCCAAACAGTGAAGAGTTTGGGGGACAGCTGATTTTGCAGACT	3940
QY	501	gttgatcccgagggtttgcagacttgcagcccaacttaccgcgttaacaaagaaagaaag	560
Db	3941	GTGGAATCCCGGGGTTTGCAGACTTGAACAGCCCACTTACCGGCTTAACCAAGAAAGG	4000
QY	561	gggattctccttgcgctcctgtagcaccagaaggtcatttgaatcacaagagccctgct	620
Db	4001	GGAAATTCCTCTGGGCTCTGAGCACCAAGGCAATTTGATGCTATCAAAAAGGCCCTGCT	4060
QY	621	gagcgacactgcttgcgacctccctgacgttaactaaaccccttacccttatttgcagta	680
Db	4061	GAGCGCACTGCTCTGCGCTCTCCGACGTAAACCCCTTACCCCTTATGCGATGA	4120
QY	681	ggtgaaggagtagccgaggaagtttlaaccacaacccctaggaacatggaggaagctgt	740
Db	4121	GGGTAAAGGAGTAGCCGAGGAGTTTAAACCAACCTAGACACTGAGAGAGACTGT	4180
QY	741	tgctcaactgtcacaagagcttgcctgtagcagatggttgcgcgtatgcttgaagc	800
Db	4181	TGCTCACTGTCAAGAAAGCTTGAATCCGTGAGCAGTGCTGGCCGTATGCTGAGAGGC	4240
QY	801	tatcgcaactgttgcacatactgtagcaggaagcgtgacaacttgaacttggagaagaat	860
Db	4241	TATCGCACTGTGCGCACTACTGCTCAAGAGCGCTGAGCAAAATGTGCTTGGGAAGATAT	4300
QY	861	aactgttaatagcccccaatgcatctggaagaacatglttcgcgcagcccccaacagatgag	920
Db	4301	AACGTAAATAGCCCCCATGATGAGAACATGTTGCGGACGCCCAACAGATGAGAT	4360
QY	921	gaccaagcccgcatgagcccaacttcaaaagcctgcttctcaagagaaggtcaacttgcg	980
Db	4361	GACCAAGCCCGCATGAGCCCACTATCAAAAGCTGCTTCTCACAGAGAGGGTCACTTCCG	4420
QY	981	tccaccagccgcttcccaacttgcacacttctgacctgaagagaactgataagcagtgac	1040
Db	4421	TCCACAGCCCGCTTCAACCTGCGCACTTCTGCTGAAAGAGACTGATGACCAAGTAC	4480
QY	1041	tcatgattgcataactatcttgaaggaacttgggttgcgaaggaacttaccataaat	1100
Db	4481	TCAATGATTGCATCAACTATTTGATTGAGAGAGATGGGGTCCGCAAGACCTTACAGACAT	4540
QY	1101	accgctgaacttgaagaagtgttaacactggttcaacttgcaggaaggaagtatgttgaag	1160
Db	4541	ACCGCTGACTGAGAAAGTGTAACTGCTGTTCACTGACGGAAGAGCTATGTGTGGAAG	4600
QY	1161	taagaagatgcttgcgagcgacgttgcacggaacccgacagatcttgcgcagcaagcct	1220
Db	4601	TAAAGATGCTGAGACGCGCGTGTGACGCGGACCCGCAAGATGTGGCCACAGACCT	4660
QY	1221	ggcggaggaagacttgcgaaagagctgaagctatgacatgacccctcaagcaagcttgcgct	1280
Db	4661	GCGGGAAGGAATTCAGCGGAAAAGGCTGACGCTCAAGCGCCCTCAAGCAAGCTTTCGCGCT	4720
QY	1281	ggcgaaggaataccaataaacttatacggacagcaagatagcttcttgcagactgca	1340
Db	4721	GGCGGAAGGAATTCATTAACATTTATACGAGACAGAGATGTGCTTGGAGACTGCA	4780
QY	1341	cgtaacaggggccaatcataaacaaggggttgccttaccctcagcagggaggaataaa	1400

```

|||||
Db 4781 CCTACACGGGGCCATCTATAACAAGGGGGTCTTACCTCAGACGGAGGGAATATAA 4840
|||
Qy 1401 gaacaagaagaattcctaagccttattagaagccttaccatttgccaaagaagctagcat 1460
|||
Db 4841 GAACAAGAGGAAATTTCTAAGCCTATTGAAGCCTTACATTGGCAAAAAGGGAGCTAT 4900
|||
Qy 1461 taatactgtccttgacacacagaagccaaagctcatalctagaagggaaccagatggc 1520
|||
Db 4901 TATACATGTCTCTGAGACATCAAGAAAGCCAAAGATCTCATATCTAGAGGGAACCGATGGC 4960
|||
Qy 1521 taaccgggttgcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagc 1580
|||
Db 4961 TAGCCGGGTTCGCAAGCAGGAGCCAGCCAGCTCTTACCTTCCCTTAATATAGAAAGGCC 5020
|||
Qy 1581 caaagcccaagaaccagaagcagtagacccctagaagcgtgcaagaagataaagaat 1640
|||
Db 5021 CAAGGCCGAGAACCCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 5080
|||
Qy 1641 agacagcttcttgagactccggaaggaagcctgctatcctcatataggaagaaatcct 1700
|||
Db 5081 AGACCAAGTCTCTGAGACATCCGAGGAGGAGCTCTATCTCATATGAGGAAGAAATCT 5140
|||
Qy 1701 gcccacaagaagaggttgaatatgtccaacagatacatcgtctaacccacttagaac 1760
|||
Db 5141 GCCCCAAGAAAGAGGTTAGATATGTCCACAGATACATGCTCACCCACCTAGAGAAC 5200
|||
Qy 1761 taacacactgcagcagttgtgcagaacatcccttataatgcttctgaggtcaccagagt 1820
|||
Db 5201 TAAACACCTGACGACGATGCTGAGAAACATCCCTTATCATGTTCTGAGAGCTACACGAGGT 5260
|||
Qy 1821 ggtgactcgtgtgttaaacattgtgtgctgcagcagcgtgtgtgttaagttaactcctcag 1880
|||
Db 5261 GGTGACTCGGTGTGTAACCATTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5320
|||
Qy 1881 aatacctccagaagaagagactaaggggaagcaccacagcgcgtcacttgggaagtgagctt 1940
|||
Db 5321 AATACCTCCGAGAAAGAGACTAAGGGGAGAACCCAGCGCTCACTGGGAATGTGACTT 5380
|||
Qy 1941 caactgagtaaacgcggcctaataacggaacaaatctcattggtttttgttagaacctt 2000
|||
Db 5381 CACTGAGGTAAACCGCGCTAAATAACGAAACAAATAATCTATGTTGTTGTGAGAACCTT 5440
|||
Qy 2001 ttaagagtaggttagagagctatcctactaagaagaagactcaacgctgtgtgtcctaaga 2060
|||
Db 5441 TTCAAGATGGGTAGAGGCTTATCTCTACTAAGAAAGAGACTTCAACCGTGTGTGCTAATAA 5500
|||
Qy 2061 aatactggagaatttttccaagatttggaaattgaaactaaggttaaggttcagacaatgg 2120
|||
Db 5501 AATACTGAGGAAATTTTTCGAGATTTCGAAATTCCTAAGGTAAATCGGGTACAGACAATGG 5560
|||
Qy 2121 tccagcttctgttgcgaagtagttagcaggaactgtgccaagatattggagattggaa 2180
|||
Db 5561 TCCAGCTTTTGTGTCGAGGATTAAGTCAAGGAGCTGGCCAAAGATATTGGGATTTGTTGAA 5620
|||
Qy 2181 actgcatgtgcacacagaccaccaagctcagaagcagtagagaagatgaataagaacat 2240
|||
Db 5621 ACTGCATTTGTGCATACAGACCCCAAGAGCTCAGAGACAGTAGAGAGATGAATGAACCAT 5680
|||
Qy 2241 taaagagaaccttaacaaattgacccacagagactgtgcatatagattgagat 2291
|||
Db 5681 TAAAGAGACCCCTTACCAAAATTTGACCAAGAGACTGGCATTAATGATTGAGT 5731
|||

```

RESULT 8
 PEN133818 8750 bp DNA VRL 14-APR-2000
 LOCUS
 DEFINITION Porcine endogenous retrovirus type C proviral gag, pol and env
 genes and LTR (class B, clone 43).
 ACCESSION AJ133818
 VERSION AJ133818.1 GI:6688949
 KEYWORDS env gene; env protein; gag gene; gag protein; pol gene; pol protein.

```

SOURCE
ORGANISM porcine endogenous retrovirus.
REFERENCE 1 (bases 1 to 8750)
AUTHORS Czauderna, F., Fischer, N., Boller, K., Kurth, R. and Tonjes, R.R.
TITLE Establishment and characterization of molecular clones of porcine
endogenous retroviruses replicating on human cells
JOURNAL J. Virol. 74 (9), 4028-4038 (2000)
MEDLINE 20219394
REFERENCE 2 (bases 1 to 8750)
AUTHORS Tonjes, R.R.
TITLE Direct Submission
SUBMITTED (04-MAY-1999) Tonjes R.R., Medical Biotechnology,
Paul-Ehrlich-Institut, Paul-Ehrlich-Strasse 51-59, Hessen, D-63225
Langer, GERMANY
FEATURES
source
location/qualifiers
1..8750
/organism="porcine endogenous retrovirus"
/proviral
/strain="type C"
/db_xref="taxon:61673"
/clone="43"
1..629
467
/feature="cap site"
632..649
/feature="g1y4"
1064..2635
/feature="gag"
1064..2635
/feature="gag"
/feature="retroviral capsid"
/product="gag protein"
join(1064..2635, 2639..6223, 6099..8072)
/feature="viral polyprotein"
/codon_start=1
/protein_id="CAB65341.1"
/db_xref="GI:6688950"
/translation="MGQVTPPLSLTLDHTEVRSRAHNLNVQVKKGPOTPCASEMP
JEDVMPSEGTENSEILAVKAIIFQNGPSSHPQDEYILTWDLADDPVAPKWLN
KRPGRPLIALGKRNKSAKVPSPRIYELIEPTWEPQVPPVPAQAVG
PSAPGAVVGGPAGATGRRGAGPERDETAIILPLTYGPMPGGLOLQVAPFS
ADLVNKNHPFSDPOLRGLVSEISMFSHPWMDCCOOLQTLFTFERERILLEA
RNNAVAGDRPTQDNEIDMGPPLTRGMDVNTAEGRESEKTYRQALVAGLRASRP
TILNAVREVMQGPENPEVLELMEARRTPTDPTSEAKASVALATGOSALDIR
KILQRLGLOEALDILVREAEKYVRRTEEEKDQREKREKREKROREKNT
KILAVVEGKSSERERDRFRIRSGPQSGNLGNRTPLDKQCAVCEKQHMARNCPK
KKNKGPVIALDEEDKGRGSDPLPEPVVALVGOVPELVDTGASHVLQPLGLK
KKNKSMWAGATGORYPMTTRTVDLAVGYTHSLVPECPVPLGRDLITKGAQI
SEQGRPEVSVNKRITVITLQLODEYRLVPOYKPDODISMLEOPQMAETAAGW
LAKQVPOVITLQKASATVSVQVRLSEAREGIMPHORLIOGILVPOSPMNTPL
LVRKRGINDVYPVODLEVNKRVDIHPYTPNPNLSALPPRNMYDILKDAF
CLRLPTSQLPFAFMROGRTGQLTWLPLQGFNSPTLEDAIRLDLANRTH
POVTLQYVDDLLAGATKODLEGTALILLESLDGLRASAKRAQICREVVLTGS
LGGORWLTLEARKTYVOIPATTTAKQVREELTAGFCRLMIPGATLAAPLPLTE
KEPESWAPRHOXTFPAIKKALISAPALAPVTPKFTLYVDERGVARGVLTQVLGQ
RRPVATLSKKIDPVAAGPVCIKALAAVALIIVKADKTLIGONTTVALPAHELVVQ
PDRMWTNARNHTYOSLTTEVYTPAPRALINPTLIPEDDEPVTDCOHLIEERG
VRKDLIDPLTEVTLTWTDSVYVEBKRNAGAAVVDGTHITASSLPBSTSQRKE
LMALDQALIAEGKSNITYDSRAFAFAHVGAIYKQGLTISAGELIKEELISL
LEALHLPLKALIIHCPGHOAKADLSRGNOMADRYAKOAOVALNLIITPKAPBR
ROYTLDMQETIKKIDQSEFTPGCTGYTSGKEILPHKGELEVOQIHRHLTHGKHQ
OLVTRSPHYVLRPLGVADSVYHCVQOLVANNPSRVPPGRKRGSHGAMVEVDFE
VPAKGNKYLIVEDPTSGVWEAVPTKKESTVYAKKILIEIPRGIKVIQSDNG
PAFVAVSOGILAKILGIDMKLHCATRGSSQGVEMAKTIKETITLKTTEGINDWA
LLEFVLEVRVNTPGOGLPYELVLGPPPLVEIASVSAVLLISQPLSEKALBEV
KORAKOLREAYSGEDLOVPHRFQVGSVYVRRRAOALFTFMKGPVLLVLTPLAV
KYVGISTWTHASHVKLAPPDGSMAETENKILRLRLVPSYNNNSPGQMHPTLSM
RHLPTGGEGKRLRPLPASTAMFLTILIPQASSKRLIDSNSHRLSLTLMILIDP
DITVYTNSTRGVARPGCTWMPDELHCLRLINPAKTSPLNIVSVYSCFPCPGTEKXC
GGSESEFCRWSCVTSNDGDKMPLISLDQRYVFSYVNSGPKYKMKLYKXKCSPSD
LDYLKISFTEKQENIOWKMGMSWGIYFTYIGGAGSTLIRLRIETGTBPVAVG

```

		PDKVLAEGPPALPEPNNLVPVOLTSLRPDITOPSPNGTGLIPNTPRNSGPVVKI GQRLFLIOGAFOAINSTDPATSSQWLSGSPYTGAMKEGKENVTKHRNOCITW GSRNKLTLFVSGKGTGKAPSPSHOHCYSYVYEOASENOYVLPGRNRMACNTGL TPCVSSVFNOSKDCVWVOLIIPRYVYPERVVLIDERYIPIRNRKREPVSLTLVMLG LGTAVGCTGTALITGPQOILEKIGELIHAAMTDLRLLESVNSLESLTSLSEVVL ONRRGLDLFLEGGGLCAALKKECCFYVDHSGALRDSKSLRERLREREREDDOGW FKGMFNRSPMMTTLTSLTALGPLVLLLTLLVGPCLINRFVAFVERISAVQIMVLROO YQGLSOGERTDL"					
mat_peptide		/gene="pol" /function="retroviral protease and polymerase" /product="Pol protein" 2639..6223					
gene		/gene="pol" /gene="env" /function="retroviral envelope" /product="Env protein" 6099..8069					
LTR		/gene="env" 8122..8750					
POLYA_signal		8643..8648					
BASE COUNT		2288 a 2214 c 2286 g 1962 t					
ORIGIN							
Query Match 49.3%; Score 1637; DB 14; Length 8750; Best Local Similarity 99.6%; Pred. No. 0; Matches 1987; Conservative 0; Mismatches 7; Indels 0; Gaps 0;							
QY	64	cttttgcttcgaaatgagatccagatccggaagaacggcgagctacatgagac	123	QY	724	ccatggaggaagactggtgctactctcaagaagcttgatcccttgagccagtgttg	783
DB	3491	CTTTTGGCTTCGAAATGAGAGATCCAGGTACGGGAAGACCGGCGACCTCACCTGAGCC	3550	DB	4151	CCATGGAGAGAACCTGTTGCTTACCTGCAAGAAGCTTGATCTCTGTAGCCAGTGTGG	4210
QY	124	cgactgcccaagaggttaagaactcccggaacatctttgcaagaagccctacagaagagac	183	QY	784	cccgatctcgaagcagctatcgaagctgtgcaactggtgcaagagagcgtgaagaattg	843
DB	3551	CGACTGCCCAAGAGGTTAAGAACCTCCCGACCACTTTTGACGAAGCCTTACACAGGAGAC	3610	DB	4211	CCCGATCTCGGAAGGCTATGCGACCTGTGGCCCTACTGGTCAAGGAGCGCTGACCAATTG	4270
QY	184	ctggccaactccagatcccaacccctcaagtgacccctcctcagtaagtgatgacgtg	243	QY	844	acttggagacgaatataactgttaatagcccccatgcatcttggaagaacatggttcgagc	903
DB	3611	CTGGCCAACCTTCAGAGATCAACACCCCTCAGGTGACCCCTCCTCCAGTGTGATGAGCTG	3670	DB	4271	ACTTTGGAGACAGATATACGTGTAATAGCCCCCATGTCATTTGGAGAACATCGTTGGCAG	4330
QY	244	ctctcgagagagcccaacaacagactgcttgaaggtacgaaggaactactgtctgaa	303	QY	904	ccccagagccgagatgagacccaagcccgcatgaaccaactatcaaaagctgtctca	963
DB	3671	CTTCTGGGGGAGCCACCAACAGAGCTGTAGAAAGGTACGAAGGCACTACTGCTGAA	3730	DB	4331	CCCCAGAGCCGATGATGAGACCAAGCCCGCATGACCCCATATCAAAACCTGCTTCTCACA	4390
QY	304	ttgtctgaacctagctacagagcctctgtctgaagaagcccgatcttgcaagaagagatg	363	QY	964	gaagaggtcaacttgcctccacacagccgctcctaacccttgcactcttgccttgaaag	1023
DB	3731	TTGTCTGACCTAGGCTACAGAGCCTCTGCTAAGAAAGCCCATGTTTCAGAGAGAGGTA	3790	DB	4391	GAGAGGCTCACTTTCGCTCCACACCGCGCTCTACACCTCTTCTGCTGAGAGAG	4450
QY	364	acataacttggaggtacagtttgaggggcgagcagatgctgacgagagcaggaagaa	423	DB	4451	actgataaacacagactgactgattgcatcaactatgtaatgagagactgggtctgc	1083
DB	3791	ACATACCTTGGGGTACAGTTTGGGGGGGAGCGATGCTGACGAGGACACGGAACAA	3850	QY	1084	aagagaccttacaagacataccgctgactgagagagtgtaacctggttcaactgacgagac	1143
QY	424	actgtagtccagatcccgcccaacacagcccaaaagtggagaggttlttggggaca	483	DB	4511	AAGACCTTACAGACATACCGCTGACTGGAAGTGTCTAACCTGTGCTACCTGACGGAAGC	4570
DB	3851	ACTGTAGTCCAGATACCGGCCCCCAACACCAACCAAGTGAAGAGATTGTTGGGACAA	3910	QY	1144	agctatgtgtgtagaagatgtcgtgagggcgcaatggttgtagcggagaccgagc	1203
QY	484	gctgagtttgcagagactgctcccggttgcagaccttgaagcagcccaacttaccg	543	DB	4571	AGCTATGTGTGGAAGTAAAGATGTGCTGGGGCGGGTGTGTGAGCGGAGCCACAGC	4630
DB	3911	GCTGAGTTTGCAGACTGTGATCCCGGGTTTGCAGCTTACAGCCCACTTACCCG	3970	QY	1204	atctgggcccagcagctgtccggaagaacttcaagcgcaaaagctgaagctatgcccctc	1263
QY	544	ctaaccaagaagaagggagatctcctgggtccttgagcacaagaagagcatcttgatc	603	DB	4631	ATCTGGGCCAGACGACCTGCGGAAGAACTTACGCCCAAAAGGCTGAGCTCATGAGCCCTC	4690
DB	3971	CTTAACCAAGAAGAGGGGAATTCCTCTGGGCTCTGAGCACCAAGACATTTGATGATGCT	4030	QY	1264	acgcaagcttgcgagctgagcgaaggaatccataaactatatacgaagcagat	1323
QY	604	atcaaaaagccctgtgagcagccctgtgacctcctctgacgtaactaaacctt	663	DB	4691	ACGCAACCTTGGCGGCTGCGCGGAAGGAATTCATTAACCTTTATACGGACACAGATAT	4750
DB	4031	ATCAAAAAGGCCCTGTGAGAGCGACCTGTCTGCGCTCTCTGACGTAACCTTAACCTTTT	4090	QY	1324	gaccttgcagctgcacagctacacgagggccatcatalaaacaaaggggtgtgtactcta	1383
QY	664	acccttattgtgagtagcgtlaagggatagcccgagagttttaaaccacaacctag	723	DB	4751	GCCTTTGGCAGCTCACACGATACACGGGGCCATCTTATTAACAAAGGGGTTGCTTACTCTA	4810
DB	4091	ACCCTTATGTGATGAGCGTAAGGAGGTAGCCGAGAGATTTTAACCCAAACCTTAGGA	4150	QY	1384	gcagggagggaaataaagaacaaagggaaattcttaagcctatagaagccttacaattg	1443

QY 331 gctaaagaagccagaattgcagagagaagtaacatacttggggtacagtttgcggg 390
DB 3847 gcttaaaagagccagatttgacagagagatgtaacttgggtacagtttgcggg 3906
QY 391 ggcagacagatgctgacagagacagagaagaaactgtagtccagataccggcccaac 450
DB 3907 gggcagacagatgctgacagagacagagaagaaactgtagtccagataccggcccaac 3966
QY 451 acagccaaacagtgagagaagtttgggagacgttgatttgcagacttgcagactcc 510
DB 3967 acagccaaacagtgagagaagtttgggagacgttgatttgcagacttgcagactcc 4026
QY 511 ggggttgcagacttgcagacccactctaccgcctaaccaagaaggggagttctcc 570
DB 4027 ggggttgcagacttgcagacccactctaccgcctaaccaagaaggggagttctcc 4086
QY 571 tgggtctctgagacacagagacatlttgatctatcaaaagggccctgtagagacac 630
DB 4087 tgggtctctgagacacagagacatlttgatctatcaaaagggccctgtagagacac 4146
QY 631 gctctggccctccctgacgttaacaaaccttacccttlatgtgtagagcgttaagg 690
DB 4147 gctctggccctccctgacgttaacaaaccttacccttlatgtgtagagcgttaagg 4206
QY 691 gtagcccaagaggtttaaaccacacccctagagacatgagagagacgttgcctacgt 750
DB 4207 gtagcccaagaggtttaaaccacacccctagagacatgagagagacgttgcctacgt 4266
QY 751 tcaaaagaagcttgatccctgtagcagagtggttggccctatctcgaagagatccga 810
DB 4267 tcaaaagaagcttgatccctgtagcagagtggttggccctatctcgaagagatccga 4326
QY 811 gtagcacaactggtgtaagaagcgttgacaatltgacttgggacagaaatalaactgtata 870
DB 4327 gtagcacaactggtgtaagaagcgttgacaatltgacttgggacagaaatalaactgtata 4386
QY 871 gcccccacatgcatltagaagaacatgcttcgagagaccccccagacagatgagacagac 930
DB 4387 gcccccacatgcatltagaagaacatgcttcgagagaccccccagacagatgagacagac 4446
QY 931 gcatltagacacacatcaaaagcctgctctcagagagaggttcaacttgcctccacagcc 990
DB 4447 gcatltagacacacacatcaaaagcctgctctcagagagaggttcaacttgcctccacagcc 4506
QY 991 gcttcaaaccttgcacacttcttgcctggaagagactgataaacagtlgactatga 1050
DB 4507 gcttcaaaccttgcacacttcttgcctggaagagactgataaacagtlgactatga 4566
QY 1051 catcaactatgatttggagagacatgggttccgagagacacttcaacataccgctgact 1110
DB 4567 catcaactatgatttggagagacatgggttccgagagacacttcaacataccgctgact 4626
QY 1111 ggaagaagtgtacacttgcctgtaacgagacagctatgtgtgaagtgaaagagatg 1170
DB 4627 ggaagaagtgtgttaacctgtgttgcctgtaacgagacagctatgtgtgaagtgaaagagatg 4686
QY 1171 gctltagggcgagtggtgtgtgcgggagacccgcagatcttgggccaagccttgcgggaag 1230
DB 4687 gctltagggcgagtggtgtgtgcgggagacccgcagatcttgggccaagccttgcgggaag 4746
QY 1231 acttcaagcgcaaaagcgtgagctcaatlggcccctcaagcaagcttggggttgcggaag 1290
DB 4747 acttcaagcgcaaaagcgtgagctcaatlggcccctcaagcaagcttggggttgcggaag 4806
QY 1291 aaatcaataacattatatacagacagcagtgatgtccttgcagactgacagcttaca 1350
DB 4807 aaatccatataacattatatacagacagcagtgatgtccttgcagactgacagcttaca 4866
QY 1351 ggcattataataaagaggggtgtgttactacagcagggagaggaagaaataaagaagaag 1410
DB 4867 ggcattataataaagaggggtgtgttactacagcagggagaggaagaaataaagaagaag 4926
QY 1411 gaaattcaagcctatagaagccttaacatttgcgaagaagcgtagctattatatacagct 1470

DB 4927 gaatttctaagcctatttataagccttacttggcccaaaagagcttgcattttracactgt 4986
QY 1471 ccttgacatcagaagaagccaaagatctacataltcagaaggaagacagatgctgacaggtt 1530
DB 4987 ccttgacatcagaagaagccaaagatccatatacagaggggaacacagatgctgacaggtt 5046
QY 1531 gccaagcagagcccaagcgttgaacttctgctataatagaagcccaagccca 1590
DB 5047 gccaagcagagcccaagcgttgaacttctgctataatagaagcccaagccca 5106
QY 1591 gaaccagaagcagatcacccctagaagacgtggaagagatgaagaagatgaccagttc 1650
DB 5107 gaaccagaagcagatcacccctagaagacgtggaagagatgaagaagatgaccagttc 5166
QY 1651 tctgagactcggagggagacgtgtacatccctataatggaagaaatccctgcccacaaa 1710
DB 5167 tctgagaaactccggagggagacgtgtacatccctataatggaagaaatccctgcccacaaa 5226
QY 1711 gaaggttgaatgtgtccaaagatatacgttcaaccacactagaactaaacactg 1770
DB 5227 gaaggttgaatgtgtccaaagatatacgttcaaccacactagaactaaacactg 5286
QY 1771 cagcagttgttcaagacalccocctatcatgttctgaggtcaccagagtgagctgactcg 1830
DB 5287 cagcagttgttcaagacalccocctatcatgttctgaggtcaccagagtgagctgactcg 5346
QY 1831 gtagtcaaaacatgtgtgctcctgcagcgttgatgaactccctccagaatacctca 1890
DB 5347 gtagtcaaaacatgtgtgctcctgcagcgttgatgaactccctccagaatacctca 5406
QY 1891 ggaagaagacatgaagggagagccaccagagcctacaggggaaggtgacacttgcaggtga 1950
DB 5407 ggaagaagacatgaagggagagccaccagagcctacaggggaaggtgacacttgcaggtga 5466
QY 1951 aagccggtcaataacggaacaaalatactatgttcttggtagaaccttccaagatg 2010
DB 5467 aagccggtcaataacggaacaaalatactatgttcttggtagaaccttccaagatg 5526
QY 2011 gtagaggttacttctactaagaagaagacatcaacggttgcctgaagaatacactgag 2070
DB 5527 gtagaggttacttctactaagaagaagacatcaacggttgcctgaagaatacactgag 5586
QY 2071 gaaatttccaaagatttgaataccttaaggtatgaaggtcagaagaatgltcaagcttcc 2130
DB 5587 gaaatttccaaagatttgaataccttaaggtatgaaggtcagaagaatgltcaagcttcc 5646
QY 2131 gtagcccaagtgatgaagcagacttgcagagataatgggagatgtatgtgaactgcattgt 2190
DB 5647 gtagcccaagtgatgaagcagacttgcagagataatgggagatgtatgtgaactgcattgt 5706
QY 2191 gcatltagac 2250
DB 5707 gcatltagac 5766
QY 2251 cttaaccaatltgacacacagagacttgcattatgaattgtagtgccttcccttgccttgc 2310
DB 5767 cttaaccaatltgacacacagagacttgcattatgaattgtagtgccttcccttgccttgc 5826
QY 2311 cttttagaggttgaaggaacacccctggaagatttgggtgacccctatgaattgtctac 2370
DB 5827 cttttagaggttgaaggaacacccctggaagatttgggtgacccctatgaattgtctac 5886
QY 2371 gggggaacccccccttgggagagaatltgcttgcacatagtgctgtagtgccttcc 2430
DB 5887 gggggaacccccccttgggagagaatltgcttgcacatagtgctgtagtgccttcc 5946
QY 2431 cagccttgttctcctagagcctcaagcgctcagatgtgttgaagcagcgctggaagag 2490
DB 5947 cagccttgttctcctagagcctcaagcgctcagatgtgttgaagcagcgctggaagag 6006
QY 2491 ctccgggagcctactccagagagagacttgcagatltcaacatcgcttccaagttgagat 2550

Db 3637 AAGCCAAAGATCTCATATATTAAGGAGACCAATGCTGACCGGGGTTGGCAAGCAGGACG 3696
Oy 1544 CCCAGCTGTAAACCTTCTGCTATAATAGAACGCCAACGCCAGAACCCAGAACGAC 1603
Db 3697 CCCAGGCTTTAACTTCTGCTATAATAGAACGCCAACGCCAGAACCCAGAACGAC 3756
Oy 1604 agtaacacctagaagactgagcaagataaanaagatagaccagttctctgagactcgg 1663
Db 3757 AGTACACCTTAGAAGACTGGCAAGAGATAAAAAGATAGACAGTTCTGTGACTCCGG 3816
Oy 1664 aaggagacctatatactataatggaagaaatccctgcccacaagaagaggttagaat 1723
Db 3817 AGGGGACCTGTATCTCATATATGGAAGAAATCTCCGCCAACAAAGAGGTTAGAT 3876
Oy 1724 atgtcaacaagataacgtctcaacccactagagactaaacactgagcaagttggtca 1783
Db 3877 ATGTCCAAAGATATACATGCTTAACCCACTAGGAACCTTGAACCTGCAAGCTTGGTCA 3936
Oy 1784 gaacatcccttatacatgttctgaggtctacagaggtgctgactcgg tgg tcaaacat 1843
Db 3937 GAACATCCCTTATCATGTTCTGAGAGGTACCGAGAGGTGGTGGTCAACATTT 3996
Oy 1844 gtgtgacctgccaagctgtgttaatgtatactctccaagataactccaggaagagataa 1903
Db 3997 GTGTGCTCCCTGACGCTGGTTATGCTAATCTCCAGAAATACCTCCAGAAAGAGACTAA 4056
Oy 1904 ggggaagccacccagagctactctggaagtggacttcaactgaagttaaagccgctaact 1963
Db 4057 GGGGAAGCCACCCAGGCGCTCACTGGGAAGTGGACTTGAAGGTTAAAGCCGGCTAAT 4116
Oy 1964 acggaacaataatctatgttcttctgagaacacttctcagagatggttagaggtctac 2023
Db 4117 ACGGAACAATAATCTATGTTGTTTGTAGACACTTTTCAGAGATGGGTAAAGGCTTATC 4176
Oy 2024 ctacttaagaagaagacttcaacctgtgtgcttaagaaatcactggaagaaatcttccaa 2083
Db 4177 CTACTAAAAAAGAGACTTCAACCGGTGGCTAAGAAATACTGGAGAAATTTTTCCAA 4236
Oy 2084 gattggaaatccctaaggttaaggttagaagaacatggtccagacttctgttccaggttaa 2143
Db 4237 GATTGGAAATACCTTAGGTAATAGGGTGAACATGCTCCAGCTTTCCTTCCCAAGTAA 4296
Oy 2144 gtcagagactgccaagatacttggagatctga tggaaactgcatgtgcatatacagacc 2203
Db 4297 GTCAAGGACTGGCCAAAGATATTTGGGGATTTGATTTGAAACCTCATTTGGCATACAGACCC 4356
Oy 2204 aaagctcagcaagtagaagagatgaataaacatlaaagagaccccttaaccaatga 2263
Db 4357 AAAGCTCAGACAGGTAGAGAGATGAATTAACCATTTAAAGAGAACCTTACCAAAATTTGA 4416
Oy 2264 ccacagaaactgcatatgatgtgagtgctccctgaccttctgttctttaggtga 2323
Db 4417 CCACAGAACTGCGATTATATGATGGATGCTCTCCCTCTTGTGCTTTTAAAGGATGA 4476
Oy 2324 ggaacaacccctggaacagtttggagctgacccctta tgaattgtcttaaggggagacccccc 2383
Db 4477 GGAACACCCCTGAGACAGTTTGGGCTGACCCCTTAAATTTGCTCTACGGGGAGACCCCCC 4536
Oy 2384 cgttggagaaatgaccttggacatatgtgtgtatgtgtgcttcccaagccttgttct 2443
Db 4537 CGTTGGCAGAAATTTGCTTGCACATATGCTGATGCTGCTCTTCCACAGCTTTGTTCT 4596
Oy 2444 cttaggctaaagcgtcgaagtgggttgaagcagagcgttgaagcagctccggaagcct 2503
Db 4597 CTAGGCTCAAGGCGCTGAGTGGGTGAGGCGACGAGCTGGAAGCAGCTCCGGAGGCGCT 4656
Oy 2504 actcaagagagagacttgaagttccacatcgcttccaagtltggaaltcagttatgtta 2563
Db 4657 ACTCAGGAGAGACTTTCAGATTCACATGCTTCCAAAGTTGGAGATTAGTATATSTA 4716
Oy 2564 gaccacacgtgaggaagaaacttgagactcgtgtggaaggaacctatctgtaactttga 2623

Db 4717 GACGCCACCGTGCAGGAACCTCGAGACTCGTAGAAGGAGCACTTATCTGACTTTTGA 4776
Oy 2624 ccacacaaagcgtgtgaagtcgaagaaatcccc 2658
Db 4777 CCACACCAAGCGCTGTGTAAGTTCGAAAGAAATCCCC 4811
RESULT 11
AR130474
LOCUS AR130474 7333 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 2 from patent US 6190861.
ACCESSION AR130474
VERSION AR130474.1 GI:14118799
KEYWORDS
SOURCE Unknown.
ORGANISM Unidentified.
REFERENCE 1 (bases 1 to 7333)
AUTHORS Fishman, J.A.
TITLE Molecular sequences of swine retroviruses method of using
JOURNAL Patent: US 6190861-A 2 20-FEB-2001;
FEATURES
source location/Qualifiers
1..7333
BASE COUNT 1984 a 1788 c 1885 g 1676 t
ORIGIN
Query Match 32.9%; Score 1091; DB 6; Length 7333;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1291; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Oy 1364 aaaggggtgtgttactctcagcagggaggaataaagaacaaaggaagaaatcttaagc 1423
Db 3517 AAAGGGGCTTGTCTTACTCAGCAGGGAGGGAATAAAGAAACAAGAAATTTCTAAGCC 3576
Oy 1424 tattaagaagcttaccatttccaaagagctatatacactgtccctgagcactaga 1483
Db 3577 TATTAGAAGCTTATCATTTGCCAAAAGCGTACTATATATACCTGTCTGGACATCAGA 3636
Oy 1484 aagcacaagatctcatatctagaggaacccagatgctgacccggttgcgaagcagcag 1543
Db 3637 AAGCCAAAGATCTCATATCTAGAGGAGACCAAGATGCTGACGGTGGCCAAACAGGACG 3696
Oy 1544 ccaaggtgttaaccttctgctatataagaagcccaagccccaagaccagacgac 1603
Db 3697 CCCAGGCTGTAACTTCTGCTATATTAAGAGGCCCAAGGCCAGAACCCAGACGAC 3756
Oy 1604 agtaacacctagaagactgcaagagataaanaagatagaccagttctctgagactcgg 1663
Db 3757 AGTACACCTTAGAAGACTGGCAAGAGATTAAGAAAGATAGACAGTTCTGTGACTCCGG 3816
Oy 1664 aaggagacctatatactataatggaagaaatccctgcccacaagaagaggttagaat 1723
Db 3817 AGGGGACCTGTATCTCATATATGGAAGAAATCTCCGCCAACAAAGAGGTTAGAT 3876
Oy 1724 atgtcaacaagataacgtctcaacccactagagactaaacactgagcaagttggtca 1783
Db 3877 ATGTCCAAAGATATACATGCTTAACCCACTAGGAACCTTGAACCTGCAAGCTTGGTCA 3936
Oy 1784 gaacatcccttatacatgttctgaggtctacagaggtgctgactcgg tgg tcaaacat 1843
Db 3937 GAACATCCCTTATCATGTTCTGAGGCTACCGAGATGCTGACTGGTGGTCAACATTT 3996
Oy 1844 gtgtgacctgccaagctgtgttaatgtatactctccaagatlaactccaggaagagactaa 1903
Db 3997 GTGTGCTCCCTGACGCTGGTTATGCTAATCTCCAGAAATCTCCAGAAAGAGACTAA 4056
Oy 1904 ggggaagccacccagagcgtcactctggaagtggacttcaactgaagttaaagccgctaact 1963
Db 4057 GGGGAAGCCACCCAGGCGCTCACTGGGAAGTGGACTTCACTGAAGGTTAAAGCCGGCTAAT 4116
Oy 1964 acggaacaataatctatgttcttctgagaacacttccaagatggttagaggtctac 2023

```

|||||
Db 4117 ACGGAACAATATCTATGTTGTTTGTAGACACCTTTTCAGATGGGTAGAGCTTATC 4176
|||
Qy 2024 ctactaagaagaagacttcaacgcttgctgctgaagaataactcggagaaattttccaa 2083
|||
Db 4177 CTACTAAAAGAGACTTCAACCGCTGGTAAAGAAAATATCTGAGAGAAATTTTTCAA 4236
|||
Qy 2084 gatttgaataactcaagtaataaggttcagacaaatgtccagcttctgcttccaggttaa 2143
|||
Db 4237 GATTGGAAATTCCTAAGTATAGGTTCAGACAAATGCTCAGCTTTGTTGCCAGGTAA 4296
|||
Qy 2144 gtccagagacttgcgaagataatltgggattgatttggaaacttgcatttgcatacagaccc 2203
|||
Db 4297 GTGAGGAGCTGGCCAGATATTGGGGATTGATTAAGAACTGCAATTTGTCATACAGACCC 4356
|||
Qy 2204 aaagctcagagcaagtagagagagatgaataaaccattaaagaaaccttaacaaatga 2263
|||
Db 4357 AAAGCTCAGAGAGTAAAGAGATGATTAAGAACCTTAAGAACCTTAACCAATTTGA 4416
|||
Qy 2264 ccacagaagacttgcataatgatgtgctcctcccttctgcttctttaggtga 2323
|||
Db 4417 CCACAGAGACTGGCATTAATGATTGATGAGCTTCCTGCTTGTGCTTTTAGGGTGA 4476
|||
Qy 2324 ggaacaccccttgcagagtttggctgcacccctatgaattgcttcaaggaggacccccc 2383
|||
Db 4477 GGAACACCCCTGGACAGTTTGGGCTGACCCCTTAATAATTGCTTACGGGGGACCCCCC 4536
|||
Qy 2384 cgttggcagaagaatttgccttgcacatagtgctgctgcttcccttccagccttcttct 2443
|||
Db 4537 CGTTGGCAGAAATTCCTTTTCACATAGTGTGCTGATGCTGCTTTCCAGCCTTTGTTCT 4536
|||
Qy 2444 ctaggctcaagcgctcagtgctgctgctgaagcagcgcttgaagcagctcccgaggagct 2503
|||
Db 4597 CTAGGCTCAAGGCGCTCGAGTGGGTGAGAGCAGCAGCGCTGGAACAGACTCCGGAGGCT 4656
|||
Qy 2504 actcagggagagacttgaatttcacacgtcttccaaatltggagattcagttatgtta 2553
|||
Db 4657 ACTCAGGAGAGACTTGCAGATTCACATTCCTTCCAGTTGGAGATTCACTATGTGA 4716
|||
Qy 2564 gaagcagacgctgcaggaagactcagactcgttgaagagaaactatctgacttctga 2623
|||
Db 4717 GACGCCACCGTCCAGAGAACTCCAGACTCGGTAGAAAGGACCTTATCTGCTACTTTGA 4776
|||
Qy 2624 ccacacaaacgctgtgaagtcgaagaatcccc 2658
|||
Db 4777 CCACACCAACGCTGTGAAGTCGAAGGAATCCCC 4811
|||
RESULT 12
AX052635 4402 bp DNA PAT 12-JAN-2001
LOCUS AX052635
DEFINITION Sequence 20 from Patent WO0071726.
ACCESSION AX052635
VERSION AX052635.1 GI:12226825
KEYWORDS
SOURCE
ORGANISM porcine endogenous retrovirus.
porcine endogenous retrovirus
Virus; Retrovirdae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
REFERENCE
1 (bases 1 to 4402)
AUTHORS Federpiet,M.J.
TITLE Xenoantigenic inhibition of infectious agent transmission during
neotransplantation
Patent: WO 0071726-A 20 30-NOV-2000;
JOURNAL Mayo Medical Ventures (US)
FEATURES
Location/Qualifiers
1..4402
/organism="porcine endogenous retrovirus"
BASE COUNT 1259 a 1085 c 1111 g 947 t
ORIGIN

```

```

Query Match 25.5%; Score 848; DB 6; Length 4402;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 2198; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

Qy 31 tcttcctgagatttaacccacatagccacacatttctgcttgaatggagatcca 90
|||
Db 1043 TTCTGCTTGAATTAACACCCACCTACCAACACCTTTTGGCTTGAATGAGATCCA 1102
|||
Qy 91 gttacggagaagaacggcgagctcacttggacccgactcggcgaaggttcaagaatcc 150
|||
Db 1103 GGTACGGGAAGAACCGGACGCTCACCTTGACCCGACTCCCCAGAGGTCAAGAACTCC 1162
|||
Qy 151 ccgacacatttgcgaagccctacacaggaacttgcgaacttgaagttccacaccc 210
|||
Db 1163 CCAGCATCTTTGACGAACCCCTACACAGAGACTTGGCCCACTTGAAGATCCACACCT 1222
|||
Qy 211 caagtacccctcctccagtagctgagtagaccccttctgaggagagcccaagaagac 270
|||
Db 1223 CAGGTACCCCTCTCCAGTACGATGATGACCTGCTTCTGGGGAGCCACCAACAGGAC 1282
|||
Qy 271 tgcctagaagtagaaggaacactacttgcgaatttgcgaacttgcgaacttgcgaact 330
|||
Db 1283 TGCTTGAAGGACGAGAGGAGACTACTGCTGATTTGCTGACCTAGGCTACAGAGCTCT 1342
|||
Qy 331 gctaaagaagccagatttgcagagagaggttaacatcttgggttgcaggttgcgggac 390
|||
Db 1343 GCTTAAGAAAGCCAGATTTGACAGAGAGATTAACCTTGTGGGTACAGTTTGGGGAC 1402
|||
Qy 391 gggcagcagatgctgcagagagcaggaagaacacttgcagatccagatccagcccaac 450
|||
Db 1403 GGGCAGCGATGGCTTGACGAGAGGACGAGGAGAAACCTGATCCAGATACCGGCCCAAC 1462
|||
Qy 451 acagccaagaatgagagatttctgggagacgcttgcgaacttgcgaacttgcgaact 510
|||
Db 1463 ACAGCCAAGAATGAGAGAGTTTGGGGACGCTGATTTGTCAGACTGATGATCCCG 1522
|||
Qy 511 gggcttgcagccttaagcagcccaactcactcagccttaacaaagaagaaggagattctc 570
|||
Db 1523 GGGTTTGGCGACTTACAGCCCTCCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 1582
|||
Qy 571 tgggctcccttaagcagcaggaagcatttgcatacaaaagcccttgcagcagcact 630
|||
Db 1583 TGGGCTCTGAGCAGCAGAGAGGATTTGATGATGATGATGATGATGATGATGATGATG 1642
|||
Qy 631 gctctggccctcctcctgacgtaactaaaccttacccttcttcttcttcttcttctt 690
|||
Db 1643 GCTCTGGCCCTCCCTGACGTAACCTTAACTTAACTTAACTTAACTTAACTTAACTTAA 1702
|||
Qy 691 gtagcccgagagattttaaaccacacacacacacacacacacacacacacacacacac 750
|||
Db 1703 GTAGCCCGGGGAGTTTAAACCAACTTAAGCAATGAGAGGAGACCTGTTGCTACCTG 1762
|||
Qy 751 tcaagaagcttgatcctcagcagcagtggttgcgctatgtctcgaaggtatcagcagct 810
|||
Db 1763 TCAAAAGAGCTCATCTCTGAGCAGTGTGTTGGCCCGGTATGCTGAAGGTATCGAGCT 1822
|||
Qy 811 gtggcactatgttcaagaagcgttgaacaaatgttacttggagacagaataactgttaata 870
|||
Db 1823 GTGGCCATGCTGCTCAAGAGCGCTGACAAATTAATGCTTGGAGACAGATTAATGTAATA 1882
|||
Qy 871 gccccecatgtcatttgaagaacatcgttgcgagcccccagcagcagcagcagcagcagc 930
|||
Db 1883 GCCCCCATATCGTTGGAGAACATCGTTGGCAGCCCCGACGAGGATGAGACCAACGCGC 1942
|||
Qy 931 cgcataaccacatacaagcctgccttcaacagagaggttcaacttgcctccacacagc 990
|||
Db 1943 CGCATGACCCACTATCAAAAGCTGCTTCTACAGAGAGAGGTGACGTTCCCTCCACAGCC 2002
|||
Qy 991 gcttcaaccccttgcacacttctgcctbgaagagacttgaagacaggttgcatactga 1050
|||
Db 2003 GCTCTCAACCTCGCACTTCTGCTGGAAGAGACATGATGAACCACTGACTCATGATGTC 2062
|||
Qy 1051 catcaactatgtatgagagacttgggttccggaagagaccttcaagaacataccgctgact 1110

```

Db 2063 CATCACTATTTGATTGAGGAGACTGGGGTCCGCAAGACCTTACAGACATACCGCTGACT 2122
 QY 1111 ggagaatgctaacctggttctactgacgaagaagcctatggtgtgaagttaagaagatg 1170
 Db 2123 GGAATAAGTGTAACTGTTGCTACAGCAGGAAGAGCTATGTGGTGAAGGTAAAGAGATG 2182
 QY 1171 gctgggacggcagctggtgagcgggacccgacgactggtgcccagcagcctgcccgaaga 1230
 Db 2183 GCTGGGGCGGGGTGTGGAGCGGAGCCCGACGATCTGGGCCAGCAGACCTGGCCGGAAGA 2242
 QY 1231 acttcagcgcaaaagcctgagctcatgcccctcaagcagccttgcggtgcccgaagg 1290
 Db 2243 ACTTCACACACAAAAGCGCTGAGCTCATGCGCCCTCACGCAACCTTTGGGGCTGGCCGAAG 2302
 QY 1291 aaatccataaccatttaagcagcagaagatgaccttgcgagctgacacgtaacggg 1350
 Db 2303 AAATCCATTAACATTATTAACAGACAGCAGATGCTTTGCGACTGCAACGCTAACAGGG 2362
 QY 1351 gccatcataaacaagggggttgcttactcctcagcagcagggggaataaagaacaaga 1410
 Db 2363 GCCATCTATAAGCAAAAGGGGTGCTTACCTCAGCAGCGAGGAATAAAGAACAAAGAG 2422
 QY 1411 gaaattcgaagcctatagaagccttactatgcccataaaggctagctattatacactgt 1470
 Db 2423 GAAATCTTAAGCCTATTAGAACCTTACATTTGCAAAAAGCGTAGCTATTATACACTGT 2482
 QY 1471 ccttgacatcagaagaagccaaagatctcatatctagaggggaacagatggtgcgggtt 1530
 Db 2483 CCTTGACATCAGAAAGCCAAAGATCCCATTCACAGAGGAAACAGATGCTGACCGGGTT 2542
 QY 1531 gcccaagcagcagcccgagctgtttaaactcttgcctataagaaagcccaagcccca 1590
 Db 2543 GCCAAGCAGCGAGCGCCAGGGGTGTTAACTTCTGCTATGATAGAAACACCCCAAGCCCA 2602
 QY 1591 gaaaccagcagcagctacacccctagaagactggaagagataaagaatagaaccagttc 1650
 Db 2603 GAACCCGCGAGCAGCTACACCTTAGAAGACTGCGCAAGATGATAAAAGATAGACAGTTTC 2662
 QY 1651 tctgagcctcggagggagactgctatacctcatatgaggaagaatccctgcggcccaaaa 1710
 Db 2663 TCTGAGACTCCGGGAAGGACCTGCTATACCTCAGATGCGAAGGAATTCCTGCCCCACAA 2722
 QY 1711 gaaaggttagaataatgcccacagatatacctgctcaaccactagaagacttaaacacctg 1770
 Db 2723 GAAGCGTTAGATATGTCCACAGATATCATGCTTAACCCACTAGAACTTAACACCTG 2782
 QY 1771 cagcagttggtcagaacatcccttcatatgcttctgaggtctacaggaagtgctgactcg 1830
 Db 2783 CAGCAGTTGGTCAGAAACATCTCCTTATCATGTTCTGAGGCTACAGAGAGTGCTGATTCG 2842
 QY 1831 gtgtgcaaacatggtgtgcacctgacgtgtaagtgtatcttcccaataactcca 1890
 Db 2843 GTGTGTAACACACTGTGTGCTCCCTGCACTGCTGTTAATGCTAATCTTCCAAATAACCTCCA 2902
 QY 1891 ggaagagactaaagggaagccacccagcgctcaactggaagtggaactcatctgagta 1950
 Db 2903 GGAAGAGACTTAAGGGAAGCACCAGCGGCTCAGTGGGAAGTGACCTTCACTGAGGTA 2962
 QY 1951 aagcggcctaataatcggaaacaataltcatgtgttctgtagaaccttctcaagatgg 2010
 Db 2963 AAGCGGCTTAATACGGAACCAATATCTATTGTTTGTAGACACCTTTTCAAGGATGG 3022
 QY 2011 gtaagccttactactactaagaagaagactcaacgtgtgtgctagaagaataactggag 2070
 Db 3023 GTAGAGCTTAATCTACTAAGAAAGAGACTTCAACCGTGTGGCTAAGAAATACTGGAG 3082
 QY 2071 gaaatttccaaagatttgtaataactcaagtaataaggtcagacaatagttccagcttcc 2130
 Db 3083 GAAATTTTTCACAAATTTTGAATATCTAAGGTAAATAGGTCAGAGCAATGTTCCAGCTTTC 3142
 QY 2131 gttgccaggttaagtcaggagactggtgcaagaatatgtggagatgtatgtgaaactgcatgt 2190
 |||||||

Db 3143 GTTCCCAAGTAAGTACAGGACTGGCCCAAGATATTTGGGAGATGATTGGAACATGCTATTGT 3202
 QY 2191 gcatagagccccaagctcagagcaggttagagagatgaatagaccattaaagacc 2250
 Db 3203 GCATACAGACCCCAAAAGCTCAGGACAGGTAGAGAGATGAATAGAACATTAAAGAGAC 3262
 QY 2251 cttaac 2255
 Db 3263 CTTAC 3267
 RESULT 13
 AX052636
 LOCUS AX052636 6076 bp DNA PAT 12-JAN-2001
 DEFINITION Sequence 21 from Patent WO0071726.
 ACCESSION AX052636
 VERSION AX052636.1 GI:12226826
 KEYWORDS
 SOURCE
 ORGANISM
 porcine endogenous retrovirus.
 porcine endogenous retrovirus
 Viruses; Retrovirdae; Mammalian type C
 retroviruses; 1-Mammalian type C virus group.
 1 (bases 1 to 6076)
 REFERENCES
 Federbpiel, M.J.
 Methods to inhibit infectious agent transmission during
 xenotransplantation
 Patent: WO 0071726-A 21 30-NOV-2000;
 Mayo Medical Ventures (US)
 JOURNAL
 TITLE
 LOCATION/Qualifiers
 FEATURES
 source
 1..6076
 /organism="porcine endogenous retrovirus"
 /db_xref="taxon:61673"
 BASE COUNT 1613 a 1512 c 1609 g 1342 t
 ORIGIN
 Query Match 24.0%; Score 797; DB 6; Length 6076;
 Best local Similarity 98.7%; Pred. No. 0;
 Matches 2197; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
 QY 31 tctgctgagattacacccacatagcacaacatttctgcttgaatgagatgaca 90
 Db 3577 TTCTGCTGAGATTTACACCCCACTAGCCACATTTTTCCTGGAATGGAGATCCA 3636
 QY 91 ggtacgggaagaacgggagagctcaacgtgagccagctgcccgaagggttcaagaatcc 150
 Db 3637 GGTACGGGAAGAACCGGAGACTCAGCTGAGACCTGAGACCTGCCAAGGGTTCAAGACTCC 3696
 QY 151 ccgaccatcttgacgaagcccttacaaggaagcctggtccaaactcaagatccaacacct 210
 Db 3697 CCGACCATCTTTGACGAAGCCCTTACACAGAGACGTGGCCAACCTTGAGATCCAACACCT 3756
 QY 211 caggtgaacctctccagtaagtgatgagctgcttctggcgaggagcccaaaagagac 270
 Db 3757 CAGGTGACCCCTCTCAGTACGTGATGACCTGCTTGGCGGAGCCACCAAAACGAGAC 3816
 QY 271 tcttgaaggtgacgaagcactactgctggaatgctgacctagagcttacaagacct 330
 Db 3817 TCGTTAAGAGGACGAAGGACCTACTGCTGGATTGTCTGACCTAGGCTACAGAGCTCT 3876
 QY 331 gctlaaagcccgagatltgacgaagagagatgaacatacttgggttacagtttgcgggac 390
 Db 3877 GCTAAGCAAGCCCAAGATTTGAGAGAGAGAGATACATTAATTGGGATGAGCTTGGCGGAC 3936
 QY 391 gggcagcagatgctgacggaagcaggaagaaactgtatgtccagatccagcccccaac 450
 Db 3937 GGGCAACGATGCTGACGAGGACAGGAGAAACTGTAGTCCAGATACCGGCCCAAC 3996
 QY 451 acagccaaacagttgagaaagcttctgggagaaagctggaatttggagactggtgactccg 510
 Db 3997 ACAGCCAAACAGTGAAGAGATTTTGGGAGACAGCTGCAATTTTGCAGACTGTGGATCCCG 4056
 QY 511 ggttgcgacttagcagcccaacttaccgctaaacaaagaaaggggagatctcc 570
 |||||||

Best Local Similarity 98.7%; Pred. No. 0;
Matches 2197; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

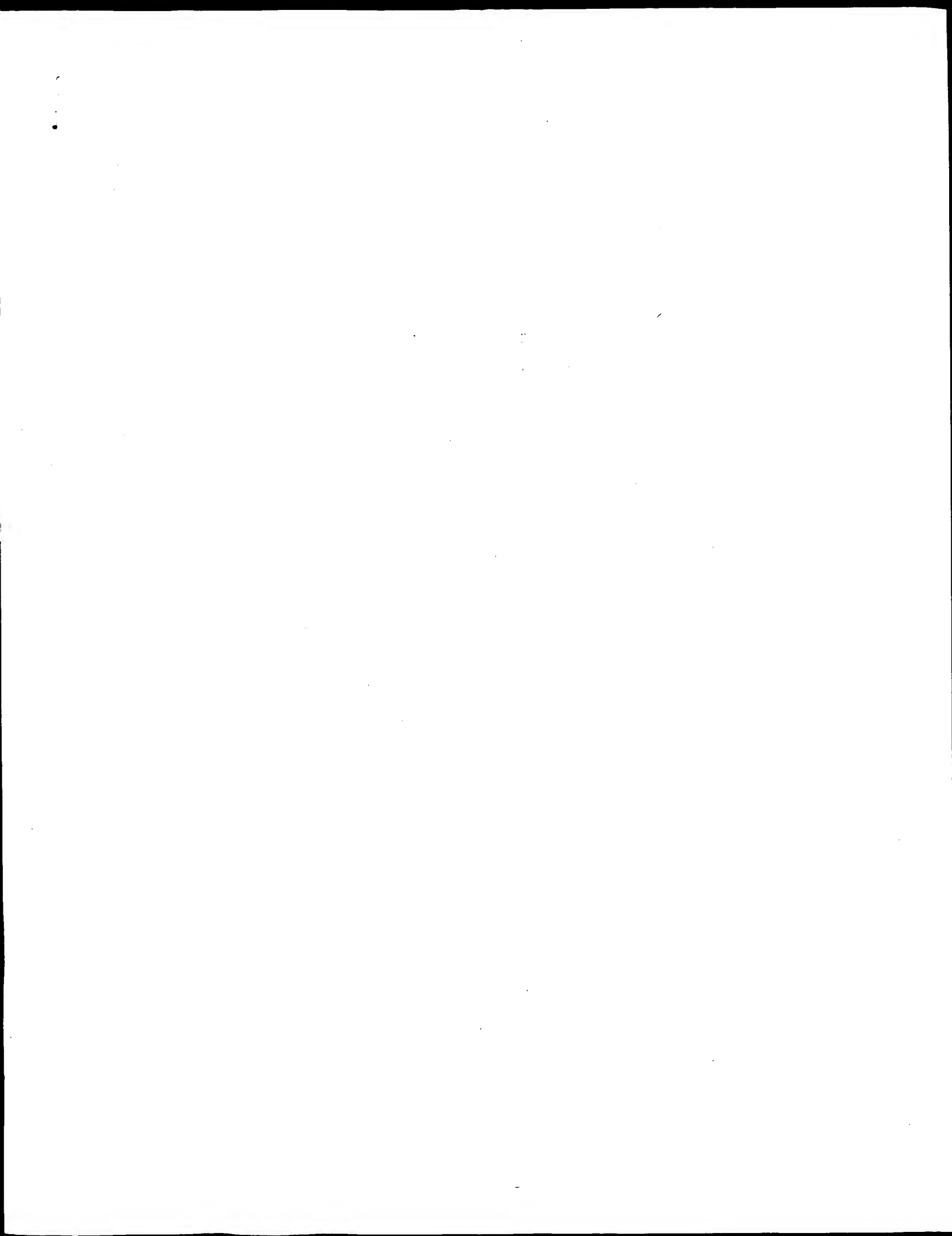
QY 31 ttctgcgtgagattacacccccctagccacacacttttgccttgatgagagatca 90
Db 3642 ttctgcgtgagattacacccccctagccacacacttttgccttgatgagagatca 3701
QY 91 ggtacgggaagaccgggagctacaccttgaccgagctgccccaaagggttaaaagacc 150
Db 3702 ggtacgggaagaccgggagctacaccttgaccgagctgccccaaagggttaaaagacc 3761
QY 151 ccgacacatttgacgaagccctacacagagagagctgccccaaatttaagatccaacacct 210
Db 3762 ccgacacatttgacgaagccctacacagagagagctgccccaaatttaagatccaacacct 3821
QY 211 caggtgaccctctccagtaagtgatgactgcttcggcgggagccaccaagaagac 270
Db 3822 caggtgaccctctccagtaagtgatgactgcttcggcgggagccaccaagaagac 3881
QY 271 tgcttagaaggtacgaagacactactgctggaattgtctgacctagagctacagagctct 330
Db 3882 tgcttagaaggtacgaagacactactgctggaattgtctgacctagagctacagagctct 3941
QY 331 gctaaagaagccagatttgacgaagagaggttaacatacttgggggtacagtttgcgggc 390
Db 3942 gctaaagaagccagatttgacgaagagaggttaacatacttgggggtacagtttgcgggc 4001
QY 391 gggcagcgatgctgacgaagagcccggaagaaactgagccagatcacgcccccaacc 450
Db 4002 gggcagcgatgctgacgaagagcccggaagaaactgagccagatcacgcccccaacc 4061
QY 451 acaagcaacaagltgagagagtttttgggagacagctgatttgcagactgtgactccg 510
Db 4062 acaagcaacaagltgagagagtttttgggagacagctgatttgcagactgtgactccg 4121
QY 511 gggtttgagacttaagcagccccacttaeccgctaaccaaaagaagggggacttctcc 570
Db 4122 gggtttgagacttaagcagccccacttaeccgctaaccaaaagaagggggacttctcc 4181
QY 571 tgggctctctgagaccacaaagacatltgatgtctcaaaaagggccctgctgagagacact 630
Db 4182 tgggctctctgagaccacaaagacatltgatgtctcaaaaagggccctgctgagagacact 4241
QY 631 gctctgcccctccctgacgtaactaaaccttaacctttagtggatgagcgtaagga 690
Db 4242 gctctgcccctccctgacgtaactaaacctttagtggatgagcgtaagga 4301
QY 691 gtagcccgagaggtttaaaccacacccctagagacatgagagagacactgtgctactcg 750
Db 4302 gtagcccgagaggtttaaaccacacccctagagacatgagagagacactgtgctactcg 4361
QY 751 tcaagaagacttgatcctgtagcagtggttggccglatgtctgaaagagctatccagac 810
Db 4362 tcaagaagacttgatcctgtagcagtggttggccglatgtctgaaagagctatccagac 4421
QY 811 gtggccactactgtcaagagacgctgacaaatltgacttgggacagaaataactgtaata 870
Db 4422 gtggccactactgtcaagagacgctgacaaatltgacttgggacagaaataactgtaata 4481
QY 871 gccccccatgatttgagaaatctgttcgcgagcccccaagacgaatgataccaagcc 930
Db 4482 gccccccatgatttgagaaatctgttcgcgagcccccaagacgaatgataccaagcc 4541
QY 931 cgaatgacccactatcaaaagcctgcttccacagagaggttcaacttgccttccacagcc 990
Db 4542 cgaatgacccactatcaaaagcctgcttccacagagaggttcaacttgccttccacagcc 4601
QY 991 gcttcaaacctgcacactcttctgcctgaagagagactgtaaacagtgactcatgttgc 1050
Db 4602 gcttcaaacctgcacactcttctgcctgaagagagactgtaaacagtgactcatgttgc 4661
QY 1051 catcaactattgtatgagagacttgggttcgcgaagagacttccagacataccgctgact 1110
Db 1110 catcaactattgtatgagagacttgggttcgcgaagagacttccagacataccgctgact 1110

Db 4662 catcaactattgtatgagagacttgggttcgcgaagagacttccagacataccgctgact 4721
QY 1111 ggaagaagtgcataaccgtgttcaactgacggaagaagctatgttggaaaglaagagatg 1170
Db 4722 ggaagaagtgcataaccgtgttcaactgacggaagaagctatgttggaaaglaagagatg 4781
QY 1171 gctggggcgagctggtgagcgaggaccgagatctggggccagcagctgcccgaagga 1230
Db 4782 gctggggcgagctggtgagcgaggaccgagatctggggccagcagctgcccgaagga 4841
QY 1231 acttcaagcgcaaaagcttgagctatgagcttcccaagcaagcttggcgctgcccgaagga 1290
Db 4842 acttcaagcgcaaaagcttgagctatgagcttcccaagcaagcttggcgctgcccgaagga 4901
QY 1291 aaatccataacatltatcaggaagcagagttatgcttggagactgacagactacaggg 1350
Db 4902 aaatccataacatltatcaggaagcagagttatgcttggagactgacagactacaggg 4961
QY 1351 gccaatcataaacaagggggttgccttacctgagcgaggaggaaataaagaacaaag 1410
Db 4962 gccaatcataaacaagggggttgccttacctgagcgaggaggaaataaagaacaaag 5021
QY 1411 gaaatltcaagcctatagaagccttaccatlttgccaaaaggtcagctatatacactgt 1470
Db 5022 gaaatltcaagcctatagaagccttaccatlttgccaaaaggtcagctatatacactgt 5081
QY 1471 cctggagatcaagaagaagcctcaatatactatagagggaacagatgctgacggggt 1530
Db 5082 cctggagatcaagaagaagcctcaatatactatagagggaacagatgctgacggggt 5141
QY 1531 gccaagcagcagcccgagctgtttaaactcttgcctataaagaagagcccaagccca 1590
Db 5142 gccaagcagcagcccgagctgtttaaactcttgcctataaagaagagcccaagccca 5201
QY 1591 gaaaccagacgagatcaacacccctagaagacttgcgaagagataaaagaatagacagctc 1650
Db 5202 gaaaccagacgagatcaacacccctagaagacttgcgaagagataaaagaatagacagctc 5261
QY 1651 tctgagactccggaggggagctgttatactcataagggaagaaatcctgcccacaa 1710
Db 5262 tctgagactccggaggggagctgttatactcataagggaagaaatcctgcccacaa 5321
QY 1711 gaaaggttgaatlatgtccaaagataacatctgtcttaaccacccctagaagaaactaaaccc 1770
Db 5322 gaaaggttgaatlatgtccaaagataacatctgtcttaaccacccctagaagaaactaaaccc 5381
QY 1771 cagcaglttggtagaacatcccttalcatalgttclgaaggtcaccagagltgctgactcg 1830
Db 5382 cagcaglttggtagaacatcccttalcatalgttclgaaggtcaccagagltgctgactcg 5441
QY 1831 gtgttcaaatatgtgtgcccctgagcagcttgggttaatgttaatccctccaagaatccca 1890
Db 5442 gtgttcaaatatgtgtgcccctgagcagcttgggttaatgttaatccctccaagaatccca 5501
QY 1891 ggaagaagactaaagggaagccacccagcgctcacttgggaagtgagacttcaatagga 1950
Db 5502 ggaagaagactaaagggaagccacccagcgctcacttgggaagtgagacttcaatagga 5561
QY 1951 aagcggcttaaalacggaacaaatatactatgtgttcttggtagacaccttccaagatgg 2010
Db 5562 aagcggcttaaalacggaacaaatatactatgtgttcttggtagacaccttccaagatgg 5621
QY 2011 gttagagcttatactacttaagaagagacttcaacgggtggtgcttaagaataactggag 2070
Db 5622 gttagagcttatactacttaagaagagacttcaacgggtggtgcttaagaataactggag 5681
QY 2071 gaaatlttccaagatlttgaatcactaaggttaataggttcaagaaatgttccacttcc 2130
Db 5682 gaaatlttccaagatlttgaatcactaaggttaataggttcaagaaatgttccacttcc 5741
QY 2131 gctggccaggtaaagtgagagacttggcccaagatlttggagatgtatgtgaactgtatgt 2190
Db 5742 gctggccaggtaaagtgagagacttggcccaagatlttggagatgtatgtgaactgtatgt 5801

QY 2191 gcatacagaccacaaagctcagagacagtagagagatagaaacattaaagagacc 2250
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 5602 GCATACAGACCCCAAACTCAGAGCAGGTAGAGAGGATGATGAACCTTAAGAGACC 5861
 QY 2251 cttaac 2255
 |||||
 Db 5862 CTTAC 5866
 |||||
 RESULT 15
 AX002802 2462 bp DNA PAT 07-SEP-2000
 LOCUS AX002802
 DEFINITION Sequence 1 from Patent WO9853104.
 ACCESSION AX002802
 VERSION AX002802.1 GI:9885131
 KEYWORDS
 SOURCE porcine endogenous retrovirus.
 ORGANISM Porcine endogenous retrovirus.
 Viruses; Retroviral viruses; Retroviridae; Mammalian type C
 retroviruses; 1-Mammalian type C virus group.
 1 (bases 1 to 2462)
 Stoye, J.P. and Weiss, R.A.
 Detection of retroviral subtypes based upon envelope specific
 sequences
 REFERENCE Patent: WO 9853104-A 1 26-NOV-1998;
 MEDICAL RES COUNCIL (GB); STOYE JONATHAN PAUL (GB); WEISS ROBIN
 ANTHONY (GB)
 JOURNAL Location/Qualifiers
 FEATURES
 source 1..2462
 /organism="porcine endogenous retrovirus"
 /db_xref="taxon:61673"
 BASE COUNT 715 a 573 c 575 g 599 t
 ORIGIN
 Query Match 23.7%; Score 786; DB 6; Length 2462;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 836; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2459 tcgagttgggtgagagcagcggtggaagagctccggagagcctactcagagagagact 2518
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 TCGAGTGGGTGAGGCGAGCGAGCTCGGAGAGCCTCGGAGGCTACTCAGAGAGAGACT 60
 QY 2519 tgcagttccacatcgcttccaaqtlggaagatlcagltcaltgtagacgacacgltgacg 2578
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 61 TGCAAGTTCCACATCGCTTCCAACTTGAGATTGAGTCTAGTAGAGCGCACCGTGCAG 120
 QY 2579 gaaacctcgagactcggtggaagagacttactctgactttgacacacacacgagctg 2638
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 121 GAAACCTCGAGACTCGGTGGAAGGAGACTTATCTCGTACTTTGACCAACAAAGGCTG 180
 QY 2639 tgaagtcgaagagatcccccactggaatccatccacgttaagcgcgcgcacacac 2698
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 181 TGAAGTGAAGGAGATCCCACTGAGATCCATGCAATCCCAAGTTAAGCCGCGCACCTC 240
 QY 2699 ccgattcggggtggaagccggaagactgaaatcccttaagcttcgctccatcgcg 2758
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 241 CCGATTGCGGGGTGGAAGGCCGAAAGACTGAAATCCCTTAAGCTTCGCTCCATCGCG 300
 QY 2759 tggttccttactctgtaataactccctcaagtaagtgaagccttgtagagagccg 2818
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 301 TGGTTCTTACTCTGTCAATTAATCTCAAGTTAATGTTAAACGCTTGTGGACAGCCG 360
 QY 2819 aactccataacccctatactcactcggttacttactgactcgcgttacaggtatattat 2878
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 361 AACTCCCATTAACCCCTTATCTCTCACTCGTTACTTACTGACTCGGTACAGGTATTAT 420
 QY 2879 attaacgacactcaagggagagctcccttgaggacctggtgagcctgaattatattgtcgc 2938
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 421 ATTAACAGCACTCAAGGGAGGCTCCCTGGGAGACTGTTGGCTGAATTATATGCTGC 480
 QY 2939 ctgcagatgaatccctcggtcctcaatgacagcgcaaccccccgatgtaactcgtgct 2998

Db 481 CTTGATCAGTATATCCCTGCTCAATGACGAGCCACACCCCGATGTACTCGTCT 540
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 2999 tacgggtttacgtttgccccagagaccccaaaatgaagaatattgtggaatccag 3058
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 541 TACGGGTTTACGTTTGGCCAGAGACCCCAATTAATGAAGATATTGGAAATCCTCAG 600
 QY 3059 gattccttgaagcaatlgagctgcataactictaaatgaggaaltggaatgagca 3118
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 601 GATTCTTTTGAAGCAATGAGCTGCAATTAATCTTAATGATGGAAATGGAGCCA 660
 QY 3119 gctctcagcaagacagagtaagttactcttctgttaacaatccaccagttatacaa 3178
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 661 GTCTCAGCAAGACAGAGTAACTTCTTTTGAACAACTTACCAAGTTATATCAA 720
 QY 3179 tttaattagccatgagagatggaagaattgcaacagcggtacaaagatgtagca 3238
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 721 TTTAATTATGCGCATGAGGATGGAAGATTGGCAACAGCGGCTACAAAGATGTACGA 780
 QY 3239 aataagcaaatagctgctgcatctgcttagaacttagtaacttaaaataagttcaact 3295
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 781 AATTAACCAATTAAGCTGTATCTGTTAGACCTAGATTTACTTAAATAAGTTTCACT 837

Search completed: February 24, 2002, 07:20:59
 Job time: 27774 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM nucleic - nucleic search, using sw model

```
Run on:      February 23, 2002, 21:01:29 : Search time 10411.9 Seconds
              (without alignments)
              5260.413 Million cell updates/sec
```

```
Title: US-09-171-553B-1
Perfect score: 3320
Sequence: 1 gaattcgcgcgcgctgcac.....aaaaaaaaaaaaaaaaaa 3320
```

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing:	Minimum Match	0%
	Maximum Match	100%

Listing first 45 summaries

Database :

1:	gb_emb1.*
2:	gb_htg.*
3:	gb_in.*
4:	gb_om.*
5:	gb_ov.*
6:	gb_pat.*
7:	gb_ph.*
8:	gb_pl.*
9:	gb_pr.*
10:	gb_ro.*
11:	gb_sls.*
12:	gb_sy.*
13:	gb_un.*
14:	gb_vl.*
15:	em_ba.*
16:	em_fun.*
17:	em_hum.*
18:	em_in.*
19:	em_om.*
20:	em_or.*
21:	em_ov.*
22:	em_pat.*
23:	em_ph.*
24:	em_pl.*
25:	em_ro.*
26:	em_sls.*
27:	em_sy.*
28:	em_un.*
29:	em_vl.*
30:	em_htg_hum.*
31:	em_htg_inv.*
32:	em_htg_rod.*
33:	em_htg_hum.*
34:	em_htg_inv.*
35:	em_htg_rod.*
36:	em_htg_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3320	100.0	3320	6	A66551	A66551 Sequence 1
2	3252.4	98.0	8849	14	PEN133817	AL133817 Porcine e
3	3227.2	97.2	8918	14	PEN233656	AU293656 Porcine e
4	3179	95.8	4402	6	AX052635	AX052635 Sequence
5	3175.8	95.7	7362	6	AX052634	AX052634 Sequence
6	2978	89.7	7873	6	AX052638	AX052638 Sequence
7	2916.8	87.9	8196	6	A66552	A66552 Sequence 2
8	2916	87.8	8209	6	A66553	A66553 Sequence 3
9	2915.2	87.8	8132	6	AR130475	AR130475 Sequence
10	2912	87.7	8132	4	AF038600	AF038600 Sus scrof
11	2876.4	86.6	8132	4	AF038600	AF038600 Sus scrof
12	2867.8	86.4	8918	14	PEN133816	AL133816 Porcine e
13	2864.6	86.3	8750	14	PEN133818	AL133818 Porcine e
14	2864.6	86.3	8763	14	PEN293657	AU293657 Porcine e
15	2857.4	86.1	7808	14	PERY17013	Y17013 Porcine eno
16	2721.4	82.0	4918	6	AX052637	AX052637 Sequence
17	2420.6	72.9	8060	6	AR130473	AR130473 Sequence
18	2415.6	72.8	6076	6	AX052636	AX052636 Sequence
19	2414	72.7	6076	6	AX052647	AX052647 Sequence
20	1771.8	53.4	7333	4	AF038601	AF038601 Sus scrof
21	1771.8	53.4	7333	6	AR130474	AR130474 Sequence
22	1175.8	35.4	8655	6	AR136169	AR136169 Sequence
23	1175.8	35.4	8655	14	AF053745	AF053745 Mus dunni
24	1168.2	35.2	154256	2	AC073295	AC073295 Mus muscu
25	1168.2	35.2	170658	2	AL596444	AL596444 Mus muscu
26	1147	34.5	274656	2	AC079538	AC079538 Mus muscu
27	1143.8	34.5	185548	10	AC005743	AC005743 Mus muscu
28	1139.4	34.3	202333	2	AC080144	AC080144 Mus muscu
29	1139.2	34.3	218938	2	AL592224	AL592224 Mus muscu
30	1135.8	34.2	196755	2	AC093350	AC093350 Mus muscu
31	1135	34.2	217034	10	AC012147	AC012147 Mus muscu
32	1132.6	34.1	65286	2	AC087800	AC087800 Mus muscu
33	1131.8	34.1	199883	2	AC073736	AC073736 Mus muscu
34	1131	34.1	3482	6	AX002804	AX002804 Sequence
35	1131	34.1	3482	6	PEREN22	Y12239 Porcine eno
36	1130.8	34.1	139317	2	AC074337	AC074337 Mus muscu
37	1129.4	34.0	257383	2	AC079494	AC079494 Mus muscu
38	1127.2	34.0	204394	10	AC079644	AC079644 Mus muscu
39	1126.4	33.9	144803	2	AC079555	AC079555 Mus muscu
40	1126.4	33.9	153567	2	AC087145	AC087145 Mus muscu
41	1126.2	33.9	188406	2	AC079560	AC079560 Mus muscu
42	1125.8	33.9	219294	2	AC015797	AC015797 Mus muscu
43	1124	33.9	181577	2	AC020959	AC020959 Mus muscu
44	1117	33.6	207677	2	AL591952	AL591952 Mus muscu
45	1115.6	33.6	158455	2	AC079471	AC079471 Mus muscu

```
BASE COUNT      922 a      844 c      832 g      722 t
ORIGIN
Query Match      100.0%; Score 3320; DB 6; Length 3320;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 gaattcgccgcgcgtcgacagatgctctctctgctgagatatacaccactagccaa 60
DB      1 GAATTCGCGCGCGCTGCACAGATGCCCTTCTGCTCCGATGATTAACACCCACTAGCCAA 60
QY      61 ccaatttggcctcgaaatgagatcccaatgctcgggaagaacccgagcaactctg 120
DB      61 CCACCTTTTGGCTTGCATGAGATCCAGATACGGGAAGAACCCGGCACTCACTCG 120
QY      121 acccgactgcccagaaggtlcaagaactcccgacacatcttctgacgaagccctacag 180
DB      121 ACCCGACTGCCCAAGGGTTCAAGACTCCCGACACATCTTTGACGAAGCCCTTACACAG 180
QY      181 gacctggccaacttcagatcccaacccctcaggtgacccctcctcagtaacgtgatgac 240
DB      181 GACCTGGCCCACTTCAGATCCCAACACCCTCAGGTGACCCCTCCAGTACGTGATGAC 240
QY      241 ctgctctcgcgagagcccaaacacagagactgcttagaagaagccagatttcagagaag 300
DB      241 CTGCTTCGCGGAGGACCCCAACCAACAGACTGCTTAGAAGAAGTGAAGGACACTACGCTG 300
QY      301 gaattgtcgaactagctacagaagcctctgctaaagaagccagatttcagagaag 360
DB      301 GAATTTGCTGACCTAGGCTACAGAGCCTGCTGAAGAGGCCCAAGATTTCAGAGAGAG 360
QY      361 gtaacactacttgggtacagatttgcgggagcgagcagatgctgacgagagcaag 420
DB      361 GTAACTACTTTGGGTACAGATTTCGGGGGCGGAGGAGGAGCTGACGGGACGCAAGAA 420
QY      421 aaaactgtagtccagataccgcccacacacagccaaagaatgagagaatgtttgggg 480
DB      421 AAAACTGTAGTCCAGATACGCGGCCACACACAGCCAAACAGTGAAGAGTTTGGGG 480
QY      481 acaagctgagatttgcagactgtgatacccgagggttgcgacacttaacagcccaactac 540
DB      481 ACAGCTGAGATTTCAGACTGTGATCCCGGGGTTTGCACCTTACACAGCCCACTAC 540
QY      541 ccgctaaccaaaaagaaggggagatttctctgggctccctgagcaccagaagacattgat 600
DB      541 CCCTTAACCAAAAGGAGGGGATTCCTCTGGGCTCTGAGCACCAAGAGGATTTGAT 600
QY      601 gctatcaaaaaggccctgtagcagcagcagctctgcccctccctgacgtaactaaccc 660
DB      601 GCTATCAAAAAGGCCCTGCTGAGCGCACTGCTGCTGCCCCCTCTGACTTAACCTAAC 660
QY      661 tttaaccttatagtgataagcgttaaggagtagcccgagaagattttaaccacaacctta 720
DB      661 TTTTACCCCTTATCTGTGATAGGCTTAAGGAGTAGCCCGAGGAGTTTAAACCAAAACCTTA 720
QY      721 ggaacca tggagagagacctttgcttaacctgtaaaagaagcttgatccctgtagcaagtgt 780
DB      721 GGACCACTGAGAGACCTTTGCTTAACCTGTAAGAAAGCTTGATCTGTACCCAGTGGT 780
QY      781 tggcccgatagtctgaaaggctatcgacgtgtgacatactgtgtaagaagccttaca 840
DB      781 TGGCCCGATGTGCTGAAGGCTATCCAGGTGTGGCCATACGTGTAAGAAGCCTTACAA 840
QY      841 ttgactttgggacagaataataactgttaataagcccccatgcaattggagaacatcgttcg 900
DB      841 TTGACTTTGGGACAGAAATAATACTGTAATAGCCCCCATGCACTTGAGACAATCGTTCGG 900
QY      901 cagccccagacagatgataagcaagcccgacatgacccactatacaagcctgcttc 960
DB      901 CAGCCCCAGACAGATGATGATGACCAAGCCCGCAATGACCCACTATCAAAAGCCTGCTTCTC 960
QY      961 acagagaggttacttctgctacacagcgcgtcacaacctgcaactcttcgctgaa 1020
DB      961 ACAGAGAGGTTACTTCTGCTCACACAGCGGCTCTCAACCCCTCCACTCTTCTGCTGAA 1020

QY      1021 gagaactgataaccagatgactcatgattgcatacaactatgattgaggaagactgggac 1080
DB      1021 GAGACTGATGAACCAAGAGACTCATGATTGCTCACTCAACTATGATGAGAGACTGGGGTC 1080
QY      1081 cgaagaagacttacaacatcacgcctgactgaggaaggttcaacctggttcaactgacga 1140
DB      1081 CGCAAGACCTTACAGACATACCGCTGACTGGAGAAAGTCTAACCTGCTTCACTGACGGA 1140
QY      1141 agcaactatgttgttgaaggttaagaagatgactgctgggacgagctgtgtgacggagcc 1200
DB      1141 AGCACTATGTGTGTGAAGTGAAGATGAGATGGCTGGGGCGGAGTGTGACCGGACCC 1200
QY      1201 acgatctggccagcagcctgcccgggaagaacttcagcgcaaaagctgagctatgacc 1260
DB      1201 ACGATCTGGCCAGCAGCCTGCGGAGAGACTTCAGCCCAAAAGGCTGAGCTCATGGCC 1260
QY      1261 ctcaagcaagcttgcgctgagcaggaaggaatcacaatacatlatacgaagcagc 1320
DB      1261 CTCAAGCAAGCTTGGCGCTGAGCGCTGCGGAGGAAATCCATTAACATTTATACGACAG 1320
QY      1321 tatgctttgcgactgacacgltacacggtgacatctataacaaagggtgtgttaacc 1380
DB      1321 TATGCTTTGCGACTGACACGCTACACGGGCGCATCTATTAACAAAGGGGTTGCTTAC 1380
QY      1381 tcagcaggaagggaataaagaacaagaagggaatltcaagctcttgaagccttacc 1440
DB      1381 TCAGCAGGAGGGAATAAAGAAACAAGAGGAATTTAAAGCTTATGAAGCCCTTACAT 1440
QY      1441 ttgccaaaaggctagctatlaatacactgtccctgacatacgaagaagccaaagatctta 1500
DB      1441 TTGCCAAAAGGCTAGCTATTATATACACTGCTCTGACACTCAGAAAGCCAAAGATCTCTA 1500
QY      1501 tctagaagggaacagatgactgacgggttgcgaagcaggaagccagctgttaacct 1560
DB      1501 TCTAGAGGGAACCAAGATGCTGACCGGGTTCGCAAGCAGGACCCAGGCTTAACTT 1560
QY      1561 ctgcctataatagaagaagcccaaaagcccaaaagcccaagcagacgtataacctgaagac 1620
DB      1561 CTGCCTATATATGAAGAGCCCAAAAGCCCAAAAGCCCAACAGCACTGTAAGAGAC 1620
QY      1621 tggcagaagatlaaaaagataagaccagttctctgagactccggaggggaaacctgtatacc 1680
DB      1621 TGGCAAGAGATAAAAAATATGACCAAGTCTCTGAGACTCCGGAGGGGACTTGTATAC 1680
QY      1681 tcaatggaaggaatactctgcccacaagaagaaggttagaatatgtccaacagatacat 1740
DB      1681 TCATATGGGAAGGAATCTCTCCCAACAAAGAGGTAGAAATATGTCAAACAGATACAT 1740
QY      1741 cgttaaccacactagaagaactaaacaactggaagaagttgtcagaataatcccttaccat 1800
DB      1741 CGTTAACCACTTAGAGACTTAACACCTGAGGAGTTGTGCAAAACATCCCTTATACAT 1800
QY      1801 gttctgaaggtacaagaagatgactgactcggttgcataaacatgtgtgacctgcaactg 1860
DB      1801 GTTCTGAGGCTAACAGAGAGTGGCTGACTCGGTGTCAAAATGTGTGCTGCTGCACTG 1860
QY      1861 gtttaagttaactcttccagaatacctccagaagaagactaaagggaagcccaaccagc 1920
DB      1861 GTTATGCTATATCTTCCAGAAATCTCCAGGAAGAGACTTAAGGGGAACCCACCAAGC 1920
QY      1921 gctcactgggaagtgagacttcaactggaagtaaacggttaactaaggaacaaatactca 1980
DB      1921 GCTCACTGGGAAGTGGACTTCACTGAGGTAAACCCGGCTAAATACGGAACCAATATCTTA 1980
QY      1981 ttggtttttagacacttttcaagatggttagagactatatacctaagaagaagact 2040
DB      1981 TTGTTTTTGTAGACACTTTTCAAGATGGGTAGAGGCTTATCTTAAGAAAGAGACT 2040
QY      2041 tcaacgtgtgtgtaagaataacttgaaggaattttccaagatttgaatactaaag 2100
DB      2041 TCAACGTTGTGTGTAAGAATAACTTGAAGGAATTTTCCAAGATTGTAATACCTAAAG 2100
```

Db 2041 TCACCGTGTGGCTAGAAATACTGAGAGAAATTTTTCOAAGATTGGAAATACCTAAG 2100
 QY 2101 gtaataaggttcagacaatggtccagcttcttgcccaagtgtaagtcggagcagcgcaag 2160
 Db 2101 GTAATAGGGTCAGACAAATGGTCCAGCTTTGTTGCCAGGATGAAGTCAGGAGCGCCAG 2160
 QY 2161 atalttggagatgtatgtggaacttgcattgtgcaacagaccccaaaagctcagagagata 2220
 Db 2161 ATATTGGGGAATGTATGGAAACTGCAATTTGCAATACAGACCCCAAAAGCTCAGGACAGTA 2220
 QY 2221 gagagagatgaatagaacacataaagagaccccttaaccaaatlgcacacagagactggcaat 2280
 Db 2221 GAGAGGATGATGATGAACCAATTAAGACACCCCTTAACCAATTTGACACAGAGACTGGCAAT 2280
 QY 2281 aatgatgagatggtctccctcgcctctgtgcttctttagagtgagagacaccccttgacag 2340
 Db 2281 AATGATTTGATGGCTCTCTCTCCCTTTGCTTTTAAAGGTGAGGAAACACCCCTGAGACAG 2340
 QY 2341 ttggagctgacccccccttaagattgtctctacgaggggagaccccccgttggcagaatggcc 2400
 Db 2341 TTTGGCTGACACCCCTATGAAATGCTCTACGGGGGACCCCGCTTGGCAGAAATTTGCC 2400
 QY 2401 ttggacaatagctgtctgt 2460
 Db 2401 TTTGACATAGT 2460
 QY 2461 gaagtggatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2520
 Db 2461 GAGTGGGTGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2520
 QY 2521 caagttccacatgcttccacatgcttccacatgcttccacatgcttccacatgcttccacat 2580
 Db 2521 CAAGTTCACATGCTTCCAGTTGGAGATTGAGTCATGATGTAGAGCCACCGTCGACGAGA 2580
 QY 2581 aaactgagagctggtgag 2640
 Db 2581 AACTCGAGAGCTGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2640
 QY 2641 aaagtccaaggaatcccccactgcatcactgcatcactgcatcactgcatcactgcatcact 2700
 Db 2641 AAAGTGAAGGAATCCCGCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2700
 QY 2701 gattcgggtgagaaagcgaagaaagcgaagaaagcgaagaaagcgaagaaagcgaagaaagc 2760
 Db 2701 GATTGGGGGTGGAAGCGCAAAAGACTGAAATCCCTTAAGCTTCGCTCCATCGCGTG 2760
 QY 2761 gttccttactctgtaataactcctcaagtaatgtgtaaacgctgtgtgagagcccgaa 2820
 Db 2761 GTTCTTACTCTGTCAATACTCTCAAGTTAAATGATGAAAGCGCTGTGTGAGACGCCGAA 2820
 QY 2821 ctccataaaccttatctctcactggttacttactgactcgcgtgacaggtatataat 2880
 Db 2821 CTCCTCAATAACCTTATCTCTACCTGCTGTTACTTACTGACTCGGTACAGGATTAATAT 2880
 QY 2881 taacagcactcaagggagggctcccttggggagctggggagctggggagctggggagctgg 2940
 Db 2881 TAACAGACTCAAGGGGAGGAGCTCCCTTGGGGAGCTGGTGGGCTGAATTAATGCTGCT 2940
 QY 2941 tggatagtaatcccttgggttcaatgacagggcacaccccccagatgactacggtgctta 3000
 Db 2941 TGCATGATATCCCTGGGTCTCATATGACAGGCGCACACCCCGGATGATCTCCGTGCTTA 3000
 QY 3001 cgggtttacgttctgcccaggaaccccaataatgaagatatgtggaatccctcagga 3060
 Db 3001 CGGGTTTACGTTTGGCCAGGAGACCCCAATAATATGAAGATATTTGGGAAATCTCTAGGA 3060
 QY 3061 ttcccttgcagaagcagtgagctgatacttcaatgtatgtggaatgtggaatgtggaatgt 3120
 Db 3061 TTTCCCTTTGCAACATGGAGCTGCAATCTTAATGATGAGGAATTTGGAATTTGGCAGAT 3120
 QY 3121 ctctcagcagaagagatgagcttcttcttcttcttcttcttcttcttcttcttcttcttct 3180
 Db 3121 CTCTCAGCAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3180

QY 3181 taatratggccattgggagatggaagattggcaacagcggggtacacaaataatgtacgaa 3240
 Db 3181 TAATTATGGCCATGGAGATGGAATTTGGACACAGCGGCTACAAAGATGATGAGAA 3240
 QY 3241 taagcaataaagctgcatctgcttagacctaacttaataaagttcactcaaa 3300
 Db 3241 TAAGCAATAAAGCTGTCATTCGTTAGACCTAGATTACTTAAAAATAGTTTCTACTAAAA 3300
 QY 3301 aaaaaaaaaaaaaaaaaa 3320
 Db 3301 AAAAAAAAAAAAAAAAAA 3320
 RESULT 2
 PEN13817
 LOCUS PEN13817 8849 bp DNA VRL 07-JAN-2000
 DEFINITION porcine endogenous retrovirus type C proviral gag, pol and env genes and LTR (class A, clone 42).
 ACCESSION AJ133817
 VERSION AJ133817.1 GI:6688947
 KEYWORDS env gene; env protein; gag gene; gag protein; pol gene; pol protein.
 SOURCE porcine endogenous retrovirus.
 ORGANISM Viruses; Retroid viruses; Retroviridae; Mammalian type C retroviruses; 1-Mammalian type C virus group.
 REFERENCE 1 (bases 1 to 8849) Toenjes R.R., Boeller K., Krach U., Kurth R. and Toenjes R.R.
 TITLE Molecular Characterization of Human-tropic and Replication-competent Porcine Endogenous Retroviruses
 JOURNAL 2 (bases 1 to 8849)
 AUTHORS Toenjes R.R.
 DIRECT SUBMISSION Submitted (04-MAY-1999) Toenjes R.R., Medical Biotechnology, Paul-Ehrlich-Institut, Paul-Ehrlich-Strasse 51-59, Hessen, D-63225 Langen, GERMANY
 FEATURES
 source location/qualifiers
 1..8849
 /organism="porcine endogenous retrovirus"
 /proviral
 /strain="type C"
 /db_xref="taxon:61673"
 /clone="42"
 /note="grown in PK15 (ATCC No. CCL-33) cell line"
 1..668
 506
 /note="cap site"
 671..688
 /note="gly4"
 1115..2686
 /gene="gag"
 1115..2686
 /gene="gag"
 /function="retroviral capsid"
 /product="Gag protein"
 join(1115..2686,2690..6271,6150..8132)
 /note="viral polypeptide"
 /codon_start=1
 /protein_id="CA65340.1"
 /db_xref="GI:6688948"
 /translation="MGCTVTPPLSLTDHTEVRSRAHNSVYQKGPQOTFCASEMP
 /TFDVGWSEGTENSEITLAVKALIFGTGPGSHDQEPYILTWODLLEDPVWVWMLN
 KPRKGPRIILALEGNKHSAEKVEPSRIYPEIEPTWEPQVPBPYPAGAVRG
 PSAPGAPVVEGPAAGTRSRGATPERTDEIALILPTKYGPMPGOLPIQWPPS
 ADLYNMTNHPPESEDPQRLTGLVESIMFSHOPWIDCOOLQTLPTTERREILAE
 KKNVPGADGPTOLONEIDMGPLTRGMDVYNAAGGESLKYRQALVGRASRP
 TNLAKVEWOGPNRPSVYLERLMEAFRFPTPDPSERKASVALATIGSALDIR
 KKLQRLGLEALRLDVRAEKVIYRRETEERKEKEREKEREKEREKEREKEREK
 KTLAAVEGSSSRERDFKIRSGPQSONLGNRPPLDKDCAYCEKSHVARNCPK
 KGNKGPVALLEEDKDRSGSDLPBRPVYLKVEGQVPEFLVDTGAHSHVLLPLGLK

QY 1399 aagaacaagaagaaatttaagccttattagaagccttactatttgcccaaaaagcctagct 1458
 DB 4877 AAGACAAAGAGAAATTTAAAGCCTATTAGAGCCTTACATTTGGCCAAAAGCCTAGCT 4936
 QY 1459 attatacaatctcttgacatacagaagaacaaagatcatactatagaggaagaccagatg 1518
 DB 4937 ATTATACAGTCTCTGAGACATCGAAGAACCAAGATCTCATATCTAGAGGAAACGATG 4996
 QY 1519 gctgacccggttgcaagaagcagccagcgtctgttaacctctgctctataagaaagc 1578
 DB 4997 GCTGACCCGGGTTCCAAAGCAGGCGCAGGCTTAACTTCTGCTTAAATAGAAAGC 5056
 QY 1579 cccaagccccaggaacccagagacagatcacacctagaagactgtgaaagataaataaag 1638
 DB 5057 CCCAAAGCCCCAGAACCCAGACACAGTACCCCTTAGAAGACTGCGAAGAGATAAAAAG 5116
 QY 1639 atagaacagttctcttgagactccggaagggagcctgtataacctataatggaagaaatc 1698
 DB 5117 ATGAGACAGTCTCTGAGACTCCGAGGGGAGCTGCTATACCTCATATGGGAAGAAATC 5176
 QY 1699 ctgcccacaagaaggggttagaataatgtccaaacagataatgtttaaaccacctaaga 1758
 DB 5177 CTGCCCCACAAAGAGGGTTAGATATGTCCAAACAGATACATGCTTCAACCCCTAGGA 5236
 QY 1759 actaaacacactgacagagttgtgcagaacataccctatcatgttcttgaggtaccaga 1818
 DB 5337 ACTAAACACCTGACAGCAGTTGGTACAGACATCCCTTATATGTCTGAGAGCTACCAAGA 5296
 QY 1819 gtggtctactcgtgtgtcaacaatgtgtccctgcccagctgtgttaagttaactctcc 1878
 DB 5297 GTGGCTGACTCGTGGTCAAAACATGTGTGCTCCCTGCGCAGCTGGTAAATGCTATCCCTCC 5356
 QY 1879 agaatacctccagaagaagaaactaaagggaggaacccacagcgtctactgtggaagtgagc 1938
 DB 5357 AGAATACCTCCAGAAAGAACTAAAGGAGGCCACCCAGGCTCTCACTGGGAAGTGAGC 5416
 QY 1939 ttcaactaggttaaagccggttaatacagaaacaatatcatatgtgttltgttagacac 1998
 DB 5417 TTCACTAGGTAAAGCCGCTTAATACGAAACAAATATCTATGTTTGTAGACACC 5476
 QY 1999 ttctcagatggttagaggtcttactactaagaagaagaaactcaacgctgtgtgtgtaag 2058
 DB 5477 TTTTCAGATGGGTAGAGGCTTATCTTAAGAAAGAGACTTCACCCGTGGTGGCTTAA 5536
 QY 2059 aaatctactgaggaatatttccaaagtttgaaactaaggaaggaataggttagagaaat 2118
 DB 5537 AAAATCTGAGGAAATTTTTCGAGATTGGAAATGGAATACCTAAGGTAATCGGGTCAAGCAAT 5596
 QY 2119 ggtccagccttctgtgcccaggttaagcaggaactgtgccaagatatgtgggatgtg 2178
 DB 5597 GGTCCAGCTTCTGTGGCCAGGTAAAGTACAGGACTGGCCAAAGATATGGGGATTTGATTGG 5656
 QY 2179 aaactgcatgtgcatacaagccccaaagctcaagacaggttagaggaatgaaatagaac 2238
 DB 5657 AAACGTGATGTGATACAGACCCCAAAAGCTCAGACAGAGTAGAGATGAATAGAAC 5716
 QY 2239 attaaagagaccccttaccaaatltgacacagagactgtgacttaatgtttgagtgtctc 2298
 DB 5717 ATTAAGAGACCCCTTACCAAAATTAACCAACAGAGACTGGCATTAATGATGGATGGCTCTC 5776
 QY 2299 ctgccccttgtgctttttagaggttagaagaacaccccttgacagtttggcctgacccctat 2358
 DB 5777 CTGCCCCCTTGTGCTTTTAAAGGTAGAGAACCCCTTGACAGTTTGGGCTGACCCCTAT 5836
 QY 2359 gaattgctctacaggggagaccccccgctgtgcagaatgtgcttgcacatagtgtgat 2418
 DB 5837 GAATGCTCTACGGGGAGACCCCGCTGTAGAAAATTCGCTCTGTAACATAGTGTGAT 5896
 QY 2419 gtgctgcttccagccttctgtctagtctgaagggcctgaggtgtgtgtgaagcaga 2478
 DB 5897 GTGGCTTTTCCAGCCCTTTTGTCTTAAGGCTCAAGGGGCTGAGTGGGTGAGCAACGA 5936
 QY 2479 gctgtgaagcagctccgggagcctactcag---aggaagactgtcgaagttccacatgc 2535

DB 5957 GCGTGAAGCAGCTCCGGAGGCGCTACTCAGAGAGAGAGACTTGCAAGTTCACATCGC 6016
 QY 2536 ttccaagttgagatltgactatgttagaacgacaggtgaggaagaactgagacagc 2595
 DB 6017 TTCCAAAGTTTGAAGATTGAGTATGTTAGAGCCACGCTGACGAAACCTGAGACTCG 6076
 QY 2596 tgaagagaccttactctgtacttltgacacacacagcgtgtgaaagtcgaagatc 2655
 DB 6077 TCGAAGGGCCCTTATCTGTAAGTTTGAACACACCAACAGCGCTGTGAAGTGAAGAAATC 6136
 QY 2656 cccactgtgacatgcatcccaagttgaagcgcgcacactcccgatltggggttgaa 2715
 DB 6137 CCCACTGTGATCCATGATCCACCTTAAGCGCGGCCACTCCCGATTCGGGGTGA 6196
 QY 2716 gccgaagaagactgaaatcccccttaagcttgcgcccaatcggtgttccctactctg 2775
 DB 6197 GCGAAGAAAGACTGAAATATCCCTTAAAGTTCGCTCCATCGGCTGCTTACTCTGTC 6256
 QY 2776 aataactcctcaagttlaagtaagcgttltgacagacccgaactcccaataacct 2835
 DB 6257 AATTAACCTCTCAAGTTAATGTAAGCGCTTGTGACAGCCGAACTCCATAAACCTT 6316
 QY 2836 atctctacactgttacttacttactgactccggtacaggtatataatatacagcactcaag 2895
 DB 6317 ATCTCAACCTGTTACTTACTGACTCCGTAACAGGTATTAATTAACAGCACTCAAG 6376
 QY 2896 ggaagctcccttgggagcctgtgtgcgaatataatgttcccttgacatgaatcc 2955
 DB 6377 GAGGCTCCCTTGGGAGCTGTGCTGATTAATATGTCTGCTTCCATCAGTAATCC 6436
 QY 2956 tgtctcaatgacagccacacaccccgatgtlactcgtgtgttcaaggtttactgtg 3015
 DB 6437 TGCTCATATGACAGGCCACACCCCGATGTACTCGCTTACGGGTTTACGTTTG 6496
 QY 3016 cccaagaccccaataatgaagaatattgtggaatctcgaagttcccttgaaga 3075
 DB 6497 CCCAGAGCCCCCAAAATTAAGAAATATGTGGAATCTCGAGATTCCTTTTGAAGCA 6556
 QY 3076 atggaactgcatacttcttaatgagatgtgaatltgacagctcctcagaagaagc 3135
 DB 6557 ATGGAAGCTCATTAATCTTAAATGATGAGGAATGGAATGGCAAGTCTCAGACAGACAG 6616
 QY 3136 agtaagttactccttltgttaacaatctcaacagttataatcaatlaattatgtgcaag 3195
 DB 6617 AGTAAGTTACTCTTGTTAACAATCCACAGTTATTAATCAATTAATTAATGAGCCATG 6676
 QY 3196 gagaatgaaagattgcaacagcgggttacaaaagaatgacgaataagcaataagctg 3255
 DB 6677 GAGATGGAAGAAATTTGCAACAGCGGTACAAAAGATGACGAATTAACGAATTAAGCTG 6736
 QY 3256 tcaatgttagacctgattacttaaaatagtttcaataaaataaaataaaataaa 3315
 DB 6737 TCATTTGTTTGAAGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6796
 QY 3316 aa 3317
 DB 6797 TA 6798

RESULT 3
 LOCUS PEN293656
 DEFINITION Porcine endogenous type C retrovirus proviral gag gene, pol gene and env gene, class A, clone 58.
 ACCESSION AJ293656
 VERSION AJ293656.1 GI:14275837
 KEYWORDS env gene; envelope; gag gene; group specific antigen; pol gene; polymerase; polypotein.
 SOURCE porcine endogenous type C retrovirus.
 ORGANISM Porcine endogenous type C retrovirus.
 REFERENCE 1 (bases 1 to 8918)

FEATURES	Location/Qualifiers
source	1. .8918

CDS	/gene="gag"
1153.	.2727

```

/gene="gag"
/codon_start=1
/product="group specific antigen"
/protein_id="CAC39616.1"
/db_xref="GI:14275838"
/translation="MSGQVTPVLSLIDHWTEYKSRAHNLSYQVKKPMQFCV
TEDVDSPEGTGFNSILLAKAIIFQTGPESHDDPEYILTMODLAEDPPMK
KPRPGRPIALIGKSNKSHAEKVPSEHIYELPEEPAMPDEQSVPPYPMQGG
SAPSGPAIVAGGPAQGTSSRRGAPERTEDIALPLRTGPIPGGOLQLOYMY
ADILNMKTNHPFSEDPORLTGLVESIMESHQTMDCOLLQDLPTEERRRIT
RKNPYGADGRPTOLONEIDMGPPLTRGCMYNTMABRSQAKIYRQALVAGLGA
TLNAKVEWVGMPNPPSVFLERLMEAFRRFTFPDTSQAKSALALATGGSAL
KKLIRQLDGLQAEHLIDVKAELKVIYKRETEEREDRKEREEREERKNKQEE
KIILAAVVEGKSNRERERDFRKIRSGPRQSNLGNRTLPDKDCAVCKEKGHWAR
KGNGLVNLALIEDKD"
/gene
/size="6309"
/length="2728"

```

Query Match	97.28;	Score 3227.2;	DB 14;	Length 8918;
Best Local Similarity	98.78;	Pred. No. 0;		
Matches 3253;	Conservative	0;	Mismatches 43;	Indels 0;
			Gaps	0;

QY	22	gagtcctctctgcccagatatacaaccccaacacagccaacactttctgcttcgaatg	81
Db	3538	gATGccTtTCTTtGcCTGgaGACTACATCCCACTAGGCAACCACTTTTCTTCGAAATG	3597
QY	82	agagatccaagtaacggaagaacccgggacagctcaactctgaccggaactggtcccaaggttc	141
Db	3598	AGGATTCACAGTGCggGAAGAAACCGGGAGCTCATTGGACTGTGACTGGCCAAAGGTTTC	3657
QY	142	aaagaactccccggacaccttttgacagaagccctacaagaaggaactggccaactcaagatc	201
Db	3658	AAAAACTCCCGGACCACTTTTGACGAGGCCCTTACACAGAGACCTGGCCAACTTCAGGATC	3717
QY	202	caaacacctcaagtgagacctcctcccaagtaagtgatgagactgctcttgccggagaacac	261
Db	3718	CAACACCCCGAGGTGACCTCTCCAGTACGTGGATGACTGCTTCTGGCGGGAGCCAC	3777
QY	262	aaacagagactgcttaagaagtaagtaaggaagacactactgctggaattcttgacctagctac	321
Db	3778	AAACAGGACTGCTTAAAGTAAGTACGAAGGCACTACGTGGAATTTCTGACTTAGCTAC	3837
QY	322	agagcctctgtcaagaagagcccaagattctcagaagaagatlaaactcttggtgagct	381
Db	3838	AGAGCCTCCGCTTAAGAGGCCAGATTCTCAGAGAGAGTAACTACTTGGGGTACAGT	3897
QY	382	ttgcggggggggcagacagatgagctgagcgagagaaagaaactgtatgctccagatacgg	441
Db	3898	TTGCGGGGGGGcAGGATGGCTGACGGAGGACGGAAGAGAAAGTGTGCAGATTACCG	3957
QY	442	gccccaacacacagccaacaagaatgagagaatttttggggagacgtcgatatttgcagactg	501
Db	3958	GCCCAACACACAGCCAAACAAAGTAGAGAGTTTGGGGACACGTGGATTTTGCGAGACTG	4017
QY	502	tggatcccggggttgggaacctttagcaaccccaactctaccggtctaaccaagaanaaagg	561
Db	4018	TGGATCCCGGGgTtTGCAcCTTACACACCCCACTCTAACCCACTTAACCAAGAAAGGAGG	4077
QY	562	ggatctcctggtgctcctgaagcaccagaaggaattgtagtctatcaaaaagagcctctg	621
Db	4078	GAATTCTCCTGGGCTCGTAGACACAGAGGCATTTGATGCTATCAAAAAGGCCCTGCTG	4137
QY	622	agcgacactgctctgagcctccctgaagtaactaaaccttaaccttaatgtgatatgag	681

D 4138 AGGCGACCTGCTGCTGGCCCTCCCTGAGCTACTAAACCTTTACCTTTATGTGGATGAG 4197
Q 682 cgtaaaggagtagcccgaggagttttaacccaaacccctaaagacatlgagagagactgtt 741
D 4198 CGTAAGGAGTAGCCCGGGAGGTTTAAACCAAACTCTAGACCATGAGACAGACCGT 4257
Q 742 gctcaactgtcacaagaagcttgatccctgttagccaagtgtgttgccgtatgtctgaagct 801
D 4258 GCGTACTGTCTAAAGAAAGCTGATCTGTAGCCAGTGGTGGCCCATATCCCTCAAGAGCT 4317
Q 802 atcgaactgtgacatactggttcaagaagcgtgacaacattgaactttggagacgaatata 861
D 4318 ATGCCACTGTGGCCATCTACTGCTCAAGAGCGCTGACAAATTGACTTTGGACAGAAATA 4377
Q 862 actgtaataagcccccaatgcatltgagagaacatgcttcgcaagcccccaagacgatgtatg 921
D 4378 ACTGTATATAGCCCCCATGGTGTGGAGACATGTTGGGAGGCCCCAGACGATGATG 4437
Q 922 accaagcccgcatgacccaactatcaaaagccttcctcaagagagagtgacttcogt 981
D 4438 ACCAAGCCCGCATGAGCCCACTATCAAAAGCCTCTCTCACAGAGAGGGTCACTGCTGCT 4497
Q 982 ccacaagccgcttcaaaccttgcacactctctgctgaagaagactatgaaccagtgact 1041
D 4498 CCACAGCCCGCTCTCAACCCCTGCCACTCTTCTGCTGAAGAGACTGATGAACCAAGTACT 4557
Q 1042 catgatgtccataactatgtatgttagagagactgggtccgcaaggaaccttacaagacata 1101
D 4558 CATGATTGGCATCACTATGATTGAGAGACACTGGGGTCCGCAAGACCTTACAGACATA 4617
Q 1102 ccgctgactgtgagaagtgtctaactcctgttcaactgacgtgaagacagctatgtgttgaaggt 1161
D 4618 CCGGTGCTGAGAAATGTTAACTGTTCACTGACGGAAGCAGCTATATGCTGGAAGGT 4677
Q 1162 aagagaatgtgtggggcgagtggtggaagcggcgaagacatcgtgggcccagcagctg 1221
D 4678 AAGAGATGCTGTGGGCGGGGTGGTGGAGCGGACCGCATGCTGGGCGCACCGCTG 4737
Q 1222 ccggaaggaacttcaagcgaacaaagctgagctcatgtgacctcaagcgaactgtgagctg 1281
D 4738 CCGGAAGGACTTCAAGACAAAAGGCTGAGCTCATGGCCCTCACGGAAGCTTTGGCGCTG 4797
Q 1282 ggcgaaggaatccatcaataacatttaacgagcagagatagcttgcgactgcacac 1341
D 4798 GCGGAAGGAAATCCATAAATTTATACAGACAGCAGATATCCCTTGCAGCTGCACAC 4857
Q 1342 gtacacggggccatctataaacaagaagggtgtgcttaacctcagcgaagggggaaataaag 1401
D 4858 GTACACGGGCGCATCTATATAAGCAAAAGGGGTGTGCTTACCTCAGCAGGAGGAAATAAAG 4917
Q 1402 aacaaagaggaatcttaagccctatagaagccttaacttgcacaaaggctagctatt 1461
D 4918 AACAAAGAGGAATTTTAAGCCATTATAGAAGCCTTACATTGGCAAAAAGGCTAGCTATT 4977
Q 1462 atacaactgtccctgagacatcaagaagccaagatctcaatctlagagggaacagatgct 1521
D 4978 ATACACTGTCTCTGACATCAAGAAAGCAAAAGATCCCATATCCAGAGGAAACAGATGGCT 5037
Q 1522 gaccgggttgccaagcagcagcagcagcgtgttaacctctgctataataaagaagccc 1581
D 5038 GACCGGGTTGCCAAGCAGGAGCCCGAGGGGTGTTAACTTCTGTCTATGTTAATAAACCC 5097
Q 1582 aaagcccccagaacccagacagtagtacaccttagaagactgtgcaagagataaanaagata 1641
D 5098 AAAGCCCCAGAAACCCGGAGACAGTACACCTTAGAAGACTGGCAAGAAATTAANAAGATA 5157
Q 1642 gaccagttctctgagactcccgagagagactgtctataactcctcatatggaagaagaatcgt 1701
D 5158 GACCAATTTCTTGAAACTCCGAGGGGACCTCCTATACCTCAGATGGGAGGAAATTCCTG 5217
Q 1702 ccccaaaaagaggttagaataatgtccaacagatatacgtctaaacccactagaaggt 1761
D 5218 CCCCAAAAGGAAGGTTAGAAATATGTCCAAACAGATATCATCTGTAAACCACTAGGAAGCT 5277

Q 1762 aaacacctgcaagcagttgttcagaaacatccccttactatgtctctgaggtaccagagtg 1821
D 5278 AAACACCTCAGACAGTGTGTCAGAAACATCCCTTATCATGTCTGAGGCTACAGAGATG 5337
Q 1822 gctgactcgtgtgttcaaacatgtgtgtccctgcagcgtgtgttaatgtactccctcaga 1881
D 5338 GCTGACTCGGTGCTCAAAATGTGTGTGCTCCAGCTGTGTTAAATGTAATCTTCCAGA 5397
Q 1882 ataacctccagaaaagagactaaagggaagcccaagcgctcactgtggaatgaaactc 1941
D 5398 ATATCTCCAGAAAGAGACTTAAGGGAGACCCACCGGGCTCCTGGAAGTGGACTTC 5457
Q 1942 actgaggtlaaagcccgctaaataacgaaacaaatactatltgttltttagaacacctt 2001
D 5458 ACTGAGGTAAACCCGGCTTAATACGAAACAATATTCATTTGTTTGTGACACCTTT 5517
Q 2002 tcaagatgggttagaggtcttactactaagaagaagacttcaaccgtgtgttaagaa 2061
D 5518 TCAGATGGGTAGAGGCTTATCTTAAGAAAGAGACTTCAACCGTGGTGGCTTAAGAA 5577
Q 2062 atactggaggaatlttccaagatttgaatataactaagtaataggttcaagacatgt 2121
D 5578 ATACTGAGGAATTTTTCAGATTGTGAATATCCTAAGTAATAGGTCAGACATAGT 5637
Q 2122 ccagcttctgtcccaagtagtcaaggaagcgtgccaagataltgaggtatgattgaa 2181
D 5638 CCAGCTTGTGTTGCCAGGTAGTCAAGGACTGGCCAGCATATTTGGGATTTGATGAAA 5697
Q 2182 ctgcattgtcatagacagaccacaagctcgaagacagtagaagagaatgaataagacatt 2241
D 5698 CTGCATTGTGCTATACAGACCCCAAGCTCAGACAGTAGAGAGATGAATAAGAACCTT 5757
Q 2242 aaagagacccttacaacattgaaccaagaagcttgatataatgtatgtatgtctcctg 2301
D 5758 AAGAGACCTTATCAAAATGTACCAAGATGAGCATTAATGATTGGATGGCTCTCTG 5817
Q 2302 ccccttctgcttcttagagtgagagacaccccttgacagcttgggtctgacccctatpaa 2361
D 5818 CCTTTGTGGCTTTTAAAGGTGAGGAACACCCCTGGACAGTTTGGCTACACCCCTATGAA 5877
Q 2362 ttgctctacgggggagcccccccgcttgacagaatltgcttgcataatgtctgattgtg 2421
D 5878 TTGCTCTAGGGGGAGCCCCCGCTTGGCAGAAATTTGCTTGGCATAATGCTGATGTG 5937
Q 2422 ctgcttccacgaccttgtctctctaggtcgaagcgctcgagtggtgtgagcagcgagcg 2481
D 5938 CTGCTTCCACACCTTTTGTCTTAAGGCTCAAGGCGCTCGAGTGGTGAAGCAGACGAG 5997
Q 2482 tgaagcagcctccggagagcctactcagagagagacttgcgaagtccacatcgcttcaa 2541
D 5998 TGGAAAGAGCTCCGGGAGAGGCTACTCAGAGAGAGACTTGCAAGTTCCATCCCTTCCA 6057
Q 2542 gttgagatctagctatgttagagccacggttgcaagaaactctgagactcgtgtgaag 2601
D 6058 GTTGAGAGTTCAAGTATATGTAGACGCCACCGTGAGGAAACCTCAGACACTCGGGAAG 6117
Q 2602 ggaacttactcgtactttgacacacaaagcgctgtgaaagtccgaaggaatcccaacc 2661
D 6118 GGACCTTATCTGTACTTTTATACACACACAGGCTGTGAAGTCAAGAGAAATCCACAC 6177
Q 2662 tgaatccatgtacccaagttlaagccgagcgccacctcccgatctcgaggttgaagaacgaa 2721
D 6178 TGGATCCATGCAATCCAGGTTAAAGCGGCGCACCTCCGATTTGGGGTGGAAAGCCGAA 6237
Q 2722 aagactgaaatccocctlaagcttcgectcacaatcggtgtgttctactactgttaataac 2781
D 6238 AAGACTGAAATCCCTTAAAGCTTCGCTCCATCGCGTGGTCTTACTGTCTCAATAAAC 6297
Q 2782 tccctcaagttaatgtlaagccttgttgagagccgaactccataaaccttatctc 2841
D 6298 TCCTCAAGTTATATGTAAACCGCTTGTGAGACGCCAGACTCCCATTAACCTTATCTCT 6357

QY 2842 cactgtgtactactactccggtacaggtattatattatcaagcactcaagggagcc 2901
|||||
Db 6358 CACCTGGTTACTTACTACTCCGGTACAGGTATTATTAACAGCACTCAAGGGAGGC 6417
2902 tcccttggggagcctgctgctgaatataatgctgctcgaatcgaatccctgctc 2961
|||||
Db 6418 TCCCTTGGGAGCCTGGTGGCTGAATATATGCTGCTGCATGATATCCCTGGTCT 6477
2962 caatgaccagagccaccccccgatgactcgtgcttgaagggttttaagttggccag 3021
|||||
Db 6478 CAATGACCAAGCCACACCCCGATGATCTGCTTACGGGTTTACGTTTGGCCAG 6537
3022 acccccaaatgaatgaatgatgctggaatccctcagatcttcccttgcgaagcaatgag 3081
|||||
Db 6538 ACCCCCAATATATGAATAATGTGGAATCTCAGAGATTTCTTTGCAAGCAATGGAG 6597
3082 ctgataactcttaataatggaatggaatggaatggaatggaatggaatggaatggaat 3141
|||||
Db 6598 CTGATTAACCTTATATGATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 6657
3142 ttaactcttggtaacatctcctcagatgataatcaatgaatgaatggaatggaatggaat 3201
|||||
Db 6658 TTACTCTTTGTTTACATCTTACAGTATTAATCAATTAATTAATGAGGATGAGAGATG 6717
3202 gaaagattggaacagcgggtacacaaagaatgaacaaatgaacaaatgaacaaatgaac 3261
|||||
Db 6718 GAAAGATTGGCAACGGGGGTACAAAAGATGACGAATTAAGCAATTAAGCTGTCAATTC 6777
QY 3262 gtttagactgattacttaataaagttcacttaaaaaaataaaaaaataaaaaa 3317
|||||
Db 6778 GTTAGACCTGATTAATTAATAAATTAAGTTCACTGAAAGAAAGCAAGAAATA 6833

RESULT 4
AX052635 4402 bp DNA PAT 12-JAN-2001
LOCUS AX052635
DEFINITION Sequence 20 from Patent WO0071726.
ACCESSION AX052635
VERSION AX052635.1 GI:12226825
KEYWORDS
SOURCE
ORGANISM
porcine endogenous retrovirus.
porcine endogenous retrovirus
viruses; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
1 (bases 1 to 4402)
AUTHORS
TITLE
METHODS to inhibit infectious agent transmission during
xenotransplantation
Patent: WO 00/1726-A 20 30-NOV-2000;
JOURNAL
Mayo Medical Ventures (US)
LOCATION/Qualifiers
FEATURES
1..4402
source /organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"

BASE COUNT 1259 a 1085 c 1111 g 947 t
ORIGIN

Query Match 95.8%; Score 3179; DB 6; Length 4402;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 3231; Conservative 0; Mismatches 65; Indels 3; Gaps 1;

QY 22 gatgacctcttcgacctgagattacacccactagcaacacacttttgcctcagatg 81
|||||
Db 1034 GATGCCCTTTTTCGCTGCAATTAACCCCACTAGCCACCACTTTTGGCTTGCATGAG 1093
QY 82 agagatccaggtacggaagaaacgggcaagctcagctggaacccagactgcccagaaggttc 141
|||||
Db 1094 AGAGATCCAGGTACGGGAAGAACCGGGCAGCTCAGCTGAGCCGAGCTGCCCAAGGTTTC 1153
QY 142 aagaactccccagacatctttgacgaagccctacacaggaactggtccaaactcagatc 201
|||||
Db 1154 AAGAATCTCCCGACATCTTTGACGAAGCCCTACAGAGACCTGGCCAACTTCAGAGATC 1213

QY 202 caaacccctcaggttacccctcctccatgacgtgatacctctctggtggagggccacc 261
|||||
Db 1214 CAACACCCCTCAAGTGAACCTCTCTCCAGTACGGGTGACCTGCTTCTGCGGAGCCACC 1273
QY 262 aacacagactgcttgaagagtaacagagcactactgctcgaatgctcgaactagctac 321
|||||
Db 1274 AAGAGAGACTGCTTGAAGAGGACGAGAGGACACTGCTGGAATGTGTGACTGAGGCTAC 1333
QY 322 agagcctctgcttaagaaagggcccaatgctcagagagaggttaacatacttgggtacagt 381
|||||
Db 1334 AGAGCCTCTGCTTAAGAAAGGCGCCAGATTGTCAGGAGAGAGGATTAATCTTGGGTACAGT 1393
QY 382 ttgctggggcggcagcagatgctgctgacagagcaggaagaacactgtagccagataccg 441
|||||
Db 1394 TTGCGGGACGGGACACGATGCTGACGAGGACGAGGAGGAGGAGGAGGAGGAGGAGGAG 1453
QY 442 gccccaacacacagcaacaaagtgagagagtttttgggagacgtggaatttgcagactg 501
|||||
Db 1454 GCCCCACCCACAGCCAAAGTGAAGAGTTTGGGAGACGCTGGATTTTGGAGACTG 1513
QY 502 tggatccgggggtttggaaccttagcagcccccactctaccgctaaccaaaagaaaggg 561
|||||
Db 1514 TGAGTCCCGGGGTTTGGCGACCTTACGACCCCTCACTTACCACTACCAAGGAGGAGGAGG 1573
QY 562 ggtctctcctgggctcctgagacacagagagagcatttgaatgctatcaaaagggcctgctg 621
|||||
Db 1574 GAATTCCTCTGGGCTCTGAGGACCAAGAGCATTTGATGCTATCAAAAAGGCCCTGCTG 1633
QY 622 agcgaactgctcgtgcccctccctcagcgaatcaaaccttacccttattgattgagtag 681
|||||
Db 1634 AGCGACCGCTGCTGGCCCTCCCTGACCTTAACCTTAACCTTTATGTGATGATGAG 1693
QY 682 cgtaaaggagtagcccgaggagttttaacccaacccagagacatgtagagagactgtt 741
|||||
Db 1694 CGTAAGGGAGTAGCCCGGGAGTTTAAACCAAACTTAGGACCAATGAGAGAGACCTGT 1733
QY 742 gctacactgcaaaagacttgatcctgtagcagtagtggccgtatgctcgaagagct 801
|||||
Db 1754 GCTTACCTGTCAAAAGACTGATCTGATGACGATGATGATGATGATGATGATGATGATGAT 1813
QY 802 atcgaactgctgacatactgctcaagaacgctgcaaaatgtaacttggagcagaata 861
|||||
Db 1814 ATCGGAGCTGTGCGCATCTGCTCAAGAGACCTGCAAAATGACTTTGGACAGCATTTA 1873
QY 862 actgtaataagcccccaatgcatgtagaagacatcgttggcagagcccccaagcagatgag 921
|||||
Db 1874 ACTGTAATAGCCCCCATGCTTGGAGAACATCGTTCCGCAACCCCGACGATGATG 1933
QY 922 accaagcccgatgacacacatcaaaagcctgcttccacagagaggttcaacttcgct 981
|||||
Db 1934 ACCAAGCCCGGATGACACCATATCAAAAGCGCTTTCACAGAGAGGATGACGCTTCCCT 1993
QY 982 ccacagacgctctcaacccctgcaactcttctgctgaagagactgtagaaccaagtagact 1041
|||||
Db 1994 CCACAGCGCGCTCTCAACCCCTGCTTCTGCTGAAGAGACTGATGAACCATGCTACT 2053
QY 1042 catgattgcatcaacttgaattgagagagactggtgctcgaagagacttaagacata 1101
|||||
Db 2054 CATGATGCGATCAACTATGATTAAGAGAGATGAGGATGAGGATGAGGATGAGGATGAG 2113
QY 1102 ccgctgacgagagagtgctaacctgctgctcagcaggaacagctacgtagtggtagaagt 1161
|||||
Db 2114 CGGCTGACGTGAGAGAGTGTAACTCGTTCACTGACGAGACAGCATATGTGTGAGAGT 2173
QY 1162 aagaagatgctggtggcgagtagtgtagacggagccgacagatctgggccaagcagcctg 1221
|||||
Db 2174 AAGAGATGAGGTGGGGGCGGTGTGTGAGCGGAGCCCGACAGATCTGGGCCAGAGGCTG 2233
QY 1222 ccggaagagacttcaagcgaagagctgagctcaatgagccctcaagcagacttggcgtg 1281
|||||
Db 2234 CCGGAAGGAACTTACGACCAAAAGGCTGAGCTCATGAGCCCTACGCAAGCTTTGGGGCTG 2293

QY 1282 gccgaagggaatcatalaacaattalacgacagcaggtatgccttgcgacltgcacac 1341
|||||
Db 2294 gccgaagggaatcatalaacaattalacgacagcaggtatgccttgcgacltgcacac 2353
QY 1342 gfacaacggggccatctataaacaagggtgtgtcttaactaagcagggagaataaag 1401
|||||
Db 2354 gfacaacggggccatctataaacaagggtgtgtcttaactaagcagggagaataaag 2413
QY 1402 aacaagaaggaaatcttaagcctatagaagccttaccatttgcgaagaagctagctat 1461
|||||
Db 2414 aacaagaaggaaatcttaagcctatagaagccttaccatttgcgaagaagcctagctat 2473
QY 1462 ataacgtctcttgcgaatcagaagaacatcatalatagaagggaacagatgct 1521
2474 ATACACTGCTCTGACATCAAGAACCCAAAGATCCCATATCCAGAGGAAACCATGGCT 2533
Db 1522 gaccgggttgcgaagcagcagccagcgtttaaactctgcctataatagaagcct 1561
|||||
Db 2534 GACCgggttgcgaagcagcagccagcgtttaaactctgcctataatagaagcct 2593
QY 1582 aaagccccaagaccagacagcagctacacccctagaagaacttgcgaagaatataaagata 1641
|||||
Db 2594 AAAGCCCAAGACCAGGACGACAGCAGTACACCTAGAGAGCTGCAAGAGATTAAGATA 2633
QY 1642 gaccagttctctgagactcaggaagggaacttgcataatagaagggaatcctg 1701
2654 GACCAGTCTCTGACATCCGGAAGGACCTGTATACCTCAGATGGGAAGGAATCTCTG 2713
Db 1702 ccccaagaagaagggttagaataatgtccaaagataatcgtctaaaccactagaact 1761
|||||
Db 2714 CCCCAGGAAGAGGGGTGAATATGTCCACAGATACATGCTTCAACCCACTTGGAACT 2773
QY 1762 aaacacttgcagcagtttgcgaacacatccctataatgtctcgaagcctacaagggtg 1821
2774 AAACACCTGCGACAGCTGTGTGAGAAACATCTCTATCATGTCTGAGGCTACCGAGAGTG 2833
Db 1822 gctgactcgttgcgaacacatgtgtcctgcagcagtggttaagtcttaactcaga 1881
2834 GCTGTGCTGTGTCGAACACATGTGTGCTGCGACAGTGTAAATGCAATCCTTCAGA 2893
QY 1882 atactcagaagaagaagcctaagggaagcaccagcgttcaacttgggaagtgaactc 1941
2894 ATACTCCAGAGAAAGAGCTTAAGGGAGGACCCAGCGCTCACTGGGAAGTGAATTC 2953
Db 1942 actgaagtaagccggttaaaatacgaagaacatactatgtgttcttgaagacactt 2001
2954 ACTGAGTAAAGCCGGCTAAATACGGAACAATATCTATGTGTTTGTAGAACCTTT 3013
QY 2002 tcaagatggttagaggtctactactaagaagaagacttcaaccgttgcgttaagaa 2061
3014 TCAGGATGGGTAGAGGCTTATCTACTAAGAAAGAGACTTCACCGTGGGTAAAGAA 3073
Db 2062 atactgagaagaatcttccagaatcttgaatagccttaaggttaaggttaagaaagct 2121
3074 ATACTGAGAGAAATTTTCCAAAGATTTGGAATACCTTAAGGTAATGAGCAAAATGCT 3133
QY 2122 ccaagcttgcgttgcgaagtaagtcaggaagcttgcgaagataatctgggatctga 2181
3134 CCAGCTTGTGTCGCCAGGTAAATCAGGAGCTGCCAAAGATTTGGGAGATTGATGAA 3193
Db 2182 ctgcatgtgtgatacagaaccccaagctcaggaagcttagaagatgaatagaacact 2241
3194 CTGATTTGTGCATACAGACCCCAAGCTCAGGACAGGTAGAGGATAGAACCAT 3253
QY 2242 aaagaagcccttaccaaatgacacagagactggaatlaataatgtagtgccttgc 2301
3254 AAAGAAGCCCTTACTAATTTGACCGGAGAGACTGGCTTATGATTTGGATAGCTCTCTG 3313
QY 2302 cccttgcgttctttagagtgaggaacaccccttgcgaagcttgcgttgcacccctaga 2361
3314 CCTTTGTGCTTTTAAAGGTAGGAACCCCTGACAGTTTGGGCTGACCCCTANTGAA 3373
Db 2362 ttgctctacagggagacccccccgttgcgaagaatctccttgcacataagtgctgtatg 2421

Db 3374 TTACTACGAGGGAGACCCCCCATTTGTGAATTTCTCCGATACATAGTGTGACGTG 3433
QY 2422 ctgcttccagccttgccttcttaagcttaagcgtcgtatggttgcgaagcagcagc 2481
3434 CTGCTTTTCCAGCTTTGTTCTTGAAGCTCAAGGCACTTGTAGTGGGAGACCAACAGCG 3493
QY 2482 tgaagcagctccggagggcctactc--agagaggaacttgaagltccaaactgcctc 2538
3494 TGGAGGCACTCCCGAGGCGCTTACTCAGAGAGGAGAGACTTGCAGATCCACATCGTTTC 3553
Db 2539 caagltggaagatcgaatcgaatgatacgcgaacgttgcgaagaacccctgcgaactcgttgc 2598
3554 CAAGTGGAGATTCAGTCACTTACGTAGACGCGCACCGTAGAGAAACCTCGAGACTCGGTG 3613
QY 2599 aagggaactatctcgttacttcttgcacacacaaagcgttgcgaagtgcgaagaatccc 2658
3614 AAGGCGCTTATACAGTATTTTGAACCAACAGGCGTGTGAAGTCGAAGGAATCTCC 3673
Db 2659 accgtgatacgaatcccaacgttgaagcagcagccacccctcagatcaggggtggaagcc 2718
3674 ACCGTGATCCATGATCCACAGCTTAAGCCGCGCCACCTCCGATTCGGGCTGGAAGCC 3733
QY 2719 gaaagagctgaaatcccttgaagcttgcctccatcgcgttgcgttcttactcgttcaat 2778
3734 GAAAGAGCTGAATATCCCTTAAAGCTTACGCTTCATTCGGGTGCTTACTCTGTCAAT 3793
Db 2779 aactcctcaagttatggttaaacgcgttgcgaacccagcagcagcactcccaataaccctatc 2838
3794 AACTCTCAAGTTAATGGAAGGCTTGTGACACCGCAACCTCCATTAACCTCTTATC 2853
QY 2839 tctcaactgttacttacttactcagcagcgttgaagatlaataatgaagcactcaagga 2898
3854 TCTCAGCTGTTACTTACTTACTTCCGCTACAGGATTAATTAATTAACAGACTCAAGGGA 3913
Db 2899 gactcccttgcggagccttgcgttgcgaatataatgtctgccttgcgaatcgaatccctgc 2958
3914 GACTCCCTTGGGAGCGCTGAGCTGAATTAATGCTGCTCGATCAGTAACTCCCTG 3973
QY 2959 tctcaatgacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3018
3974 TCTCAATGACACAGGCAACCCCGCATGATCTGCTTACGGGTTTACGTTTGGCCC 4033
Db 3019 aggaaccccaataatgaagaatataatgttgaatccctcaggttctccttgcgaagcagc 3078
4034 AGGACCCCAATATGAAGAAATATGTGGAATCTCGAGATTTCTTTTGCAGAGCAATG 4093
QY 3079 gactgataacttcttaatgataatggaatgcgaatgcgaatgcgaatgcgaatgcgaatgc 3138
4094 GAGCTCGTAACCTTAATGATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 4153
Db 3139 aagttactcttgcgttaacaactccagcagcagcagcagcagcagcagcagcagcagc 3198
4154 AAGTTACTCTTTTGTAAACATCTTACAGTTATATTAATTAATTAATTAATTAATTAAT 4213
QY 3199 atggaagaatcttgcgaacagcagcagcagcagcagcagcagcagcagcagcagcagc 3258
4214 ATGGAAGAAATGGAACAGCGGCTACAAAAGATGTACAAATTAACCAATTAAGCTGTGA 4273
Db 3259 ttcgttgaagccttgaatcttaaaataagccttcaaaataaaataaaataaaataaa 3317
4274 TTTCGTTAGACCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4332
RESULT 5
AX052634 7362 bp DNA PAT 12-JAN-2001
LOCUS AX052634 Sequence 19 from Patent WO0071726.
DEFINITION AX052634
ACCESSION AX052634
VERSION AX052634.1 GI:12226824
KEYWORDS
SOURCE porcine endogenous retrovirus.
ORGANISM porcine endogenous retrovirus

Db 5493 ATACCTCCAGGAAGAGACTAAGGGAGAACCCACCGCTCCTACTGGAGATGACATTC 5552
QY 1942 actgaggtaaagccggtaataatcgaaacaaatctctgtgttttttgaagcacttt 2001
Db 5553 ACTGAGTTAAAGCGGCTTAATACGGAAACAATATCTATGTTGTTTGAAGCACCTTT 5612
QY 2002 tcaagatggttagaggtcttctactaagaagaagactcaacgctgtgtgtctaaaga 2061
Db 5613 TCAGGATGGGTAGAGGTTATCTACTAAGAAAGAGACTTCAACCGTGGGTAAAAAA 5672
QY 2062 atactgaggaagaattttccaaagatttgaaatccaaagtaaggttcagaacatgt 2121
Db 5673 ATACTGGAGGAATTTTCCAAAGATTGGAATCTAAGTATATAGGTGACAGCAATGGT 5732
QY 2122 cgaagcttggttcccaaggaatgcaggtcgtgccaagatatgtgggttgaatga 2181
Db 5733 CCAAGCTTTCCTGCCCCAGGTAAGTCAAGGACTGCGCAAGATATGGGGATGTTGGAAA 5792
QY 2182 ctgcatgtgcatcacagaccccaagctcagagcaggtagagagatgataagaacatt 2241
Db 5793 CTGCATTTGTCATACAGAACCCCAAACTCAGGACAGGTAGAGAGATGATAGAACATT 5852
QY 2242 aaagagaaccttaccaaattgaccacagagactgcatatgattgagtgtctctgt 2301
Db 5853 AAAGAGACCCCTTACTAATTTGACCGCGAGACTGGCTTAATGATTGATAGTCTCTG 5912
QY 2302 cctcttggtgttttaggtgtgaagaaacacccctgagcaatttggtgtgccccatatga 2361
Db 5913 CCTTTGTACTTTTATGGGTTAGGAACACCCCTGGACAGTTGGGCTGACCCCTATWGA 5972
QY 2362 ttgcttcaagggagacccccccctgtgtgcagaatgtccttgcacatgtgtgtgt 2421
Db 5973 TTACTCTACGGGGAGACCCCCCATTTGTAGAAATTTGTTCCGTATAGTACTGACGTG 6032
QY 2422 ctgcttcccaagcttgttctcttaagcccaagggcgtcgtgagtgggtgagcagcgagc 2481
Db 6033 CTGCTTTCCAGGCTTTGTTCTCTTAAGGCTCAAGGCACTTGAAGTGGTGAACACAGAGCG 6092
QY 2482 tgaagcaagctccggagagcctactc---agagagagacttcaagttccacatgccttc 2538
Db 6093 TGGAGGCAACTCCGGGAGGCTTACTAGAGAGAGAGAGACTTGCAGATCCCAATCGTTTC 6152
QY 2539 caagttggaagatcagltctatglttagacgcacccgtgcagagaacctcgaagactcgltg 2598
Db 6153 CAAGTGGAGATTCAGTCTACGTTTACGCCCAACCGGTGAGGAACCTCGAGACTGGTGTG 6212
QY 2599 aagggaccttatctgtacttttgaccacacaaagcgtgtgaagtcgaagaaatcccc 2658
Db 6213 AAGGGCCCTTATCACGTAATCTTTTGAACACACAGCGCTGTGAAAGTCAAGAAATCTTC 6272
QY 2659 acctgacatgacatccacagtttaagcggcgccacccctccgattcggtgtggaagcc 2718
Db 6273 ACTGTGATCATGATCCCAAGTTAAAGCCGGCCACCTCCCGATTCGGGGTGAAGGCC 6332
QY 2719 gaaagaaactgaaatcccccttaagcttcgctccatcgcgtgtgttccctactcttcaat 2778
Db 6333 GAAAAAGACTGAAATCCCTCTTAAGCTTCGCTTCATCGGCTGCTTACTCTGTCAAT 6392
QY 2779 aactctaaagttaatgtaaaagcgtgtgtgaagcccggaactcccaataaaccctatc 2838
Db 6393 AACTCTCAAGTTAATGTAAGCGCTTGTGACAGCCCGCACTCCCAATAACCCCTTATC 6452
QY 2839 tctcacactgttactactgactcagtcagatgatttaataataacagaaactcaagggga 2898
Db 6453 TCTCACACTGTGTTACTTACGACTCCGGTACAGGTATTAATTAACACGACACTCAAGGGGA 6512
QY 2899 ggtcccttgggaacctgtgtgcgtgaattatagttctgtccttcgaatcaatccctgt 2958
Db 6513 GGTCCCTTGGGGACCTGGTGGCTGAATTAATATGCTGCTTCATCAGTAATCCCTGG 6572
QY 2959 tctcaatgacagagccacaccccgatgatactcgtgtcttaaggttttaagtttgc 3018
Db 6573 TCTCAATGACCAAGGCCACACCCCGGATGTAATCGGTGCTTACGGGTTTACGTTTGGCC 6632

QY 3019 agagcccccaataatgaagaatattgtggaatcctcaagattccttgcagaagatg 3078
Db 6633 AGGACCCCAATTAATGAAGATATTGTGAATCTCAGATTTCTTTTCAAGCAATG 6692
QY 3079 gagctgataacttactatgaggaatgtgaatgtgcagctctcctagcaagcagagt 3138
Db 6693 GAGCTGCGTAACCTTATATGATGGGAATTTGGAATGGCCATGCTCTCAGCAAGACAGT 6752
QY 3139 aagttactcttcttgaacaaatccctacagttataatcaatlaattlaattggtccatggag 3198
Db 6753 AAGTACTCTTTTGTAAACAATCTTACCAAGTATATATCAATTTATATATGAGCAATGGAG 6812
QY 3199 atggaagaattgtgcaacagcggtgtcaaaagaatgtacagaataagaacaaatgtca 3258
Db 6813 ATGGAAGAATTGTGCAACGCGGTCAAAAGATGTACGAATAATGAAGTGTGCA 6872
QY 3259 ttgcgttagaaccttagaattacttaaaataagtttcaactaaataaaataaaataaa 3317
Db 6873 TTGCTTAGACCTAGATTACTTAAATATAGTTTCTACTGAAAAAGGAACAAAGAAATA 6931

RESULT 6
AX052638 AX052638 7873 bp DNA PAT 12-JAN-2001
LOCUS
DEFINITION Sequence 23 from Patent WO0071726.
ACCESSION AX052638
VERSION AX052638.1 GI:12226828
KEYWORDS
SOURCE
ORGANISM porcine endogenous retrovirus.
porcine endogenous retrovirus
Virus; Retroid viruses; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
REFERENCE
AUTHORS Federpiel, M.J.
TITLE Methods to inhibit infectious agent transmission during
xenotransplantation
JOURNAL Patent: WO 0071726-A 23 30-NOV-2000;
Mayo Medical Ventures (US)
FEATURES
source
1. 7873
Location/Qualifiers
/organism="porcine endogenous retrovirus"
BASE COUNT 2200 a 1914 c 1961 g 1798 t
ORIGIN

Query Match 89.7%; Score 2978; DB 6; Length 7873;
Best Local Similarity 95.2%; Pred. No. 0;
Matches 3140; Conservative 0; Mismatches 70; Indels 90; Gaps 3;
QY 22 gatgcctctctgcgcgtgagatlaacaccacatagcaaccacttctgcctcgaaatg 81
Db 2447 GATGCTTTTCTGCTCGAGATTACACCCCACTAGCCACACACTTTTGGCTTGAAATG 2506
QY 82 agagatccaggttaggaagaacccggcagctcctgcgcgactgcccgaaggttc 141
Db 2507 AGAGATCCAGGTACGGGAGAAACCGGAGCTCAGCTGAGACCGACGACGCCCCAAGGGTTC 2566
QY 142 aagaactcccgagacatcttgaagaaagccctacacagggagactgcgcaactcaagatc 201
Db 2567 AAGAAGTCCCGGACACATCTTTGAGAGAGCCCTACACAGAGACCTGGGCAACTTACGATC 2626
QY 202 caaacacctgaagtgaacctctcctcagtaagttgaagactgcttgcgaggaagcacc 261
Db 2627 CAACACCTCAGGTGACCTCTCCTCAGTACCTGATGACTGCTTTCGGGGGAGCCACC 2686
QY 262 aaacagagactcttagaaggttagaaggaagcactactgctgaatgtgtgactaggttac 321
Db 2687 AAACAGGACTGCTTAGAAGACGACGACATTAATGCTGGAATGTGTGACTGAGCTTAC 2746
QY 322 agagcctctgctaaagagcccaagatttgacaggaaggttaacatacttgggtacagt 381

|||||
Db 3916 GCTCCACGACCGCTCTCAACCTGCGACTTTCGCCGAGAGACATGATTAACCGAGTG 3975
OY 1039 actcatgatctgcatcaatcatatgatitgagagactgggtccgcaagaccttcaagc 1098
|||||
Db 3976 ACTCATGATTTGCCATCACTATTGATGAGAGACTGGGGTCCGCAAGACCTTTACAGAC 4035
OY 1099 ataccgtactgtagaagtgctaaacctggttcaactgacggaagcagctatgtgtgaa 1158
|||||
Db 4036 ATACCCCTGACTGAGAAATGCTTAACCTGTTCACTGACGAGACGATGATGTGTGANA 4095
OY 1159 gctaaagagatgctggtggcgagtggtgtgagccggaccgcaagcttggcgagc 1218
|||||
Db 4096 GGTAAAGGATGCTGGGGGAGGTGGGAGCGGAGCCCGCAGCATCTGGGCGACGACG 4155
OY 1219 ctgcccgaaggaactcagcgcaaaagctgagctcatgagccctcaagcaagcttggc 1278
|||||
Db 4156 CTGCGGAGAGAACTTACACCGCAAAAGGTGAGCTCATGCGCCCTCACCAACCTTTGGG 4215
OY 1279 ctgcccgaaggaagatccaataacattatagcaagcaggtatgcttctgagctgca 1338
|||||
Db 4216 CTGGCGGAAGGGAATCCATAACATTATACGACAGAGATGCTCTTGCCGACTGCA 4275
OY 1339 cacttacacggggccatcctataaacaagggtgtgtcttactcaagcgaggagaaata 1398
|||||
Db 4276 CACGTACACGGGGCATCTATTAACAAGGGGTTGCTTACTCAGCAGGAGGGAAYTA 4335
OY 1399 aagaaacaaagaggaatcttaagcctatagaagccttcaacttgcacaaagctagct 1458
|||||
Db 4336 AAGAACAAAGAGAGAAATCTTAAGCCTATTAGAGCCTTACATTTCGCAAAAAGGTACT 4395
OY 1459 attatacactgtcctcgagacatcagaagccaaagatcctacatctagaaggaaacagatg 1518
|||||
Db 4396 ATTATACACTGTCTCGGACATCCGAAAGCCAAAGATCTCATATCTAGAGGGAACGATG 4455
OY 1519 gctgacgggtgttgcacaaagcagccagcgtgttcaaccttgcctataatagaagc 1578
|||||
Db 4456 GCTGACCGGGTTGCCACAGCAGGACGCCAGGCTGTTAACTTCTGCTTATATATGAACG 4515
OY 1579 cccaaagcccaagaaacccagacgacagcagccttagaagactggaagaaatataaag 1638
|||||
Db 4516 CCCAAAGCCCGAGAACCCAGACGACAGTACACCTAGAAGACTGGCAAGACATTAATAAG 4575
OY 1639 atagacaggttctctgaaaccccgagagagcgtctactacataatggaagaaatc 1698
|||||
Db 4576 ATAGACCAAGTCTCTGAACTCCGAGGGGACCTGCTATACCTATATAGGAAAGAAATC 4635
OY 1699 ctgcccacaaagaaaggttagaatatgtccaagatatactgttcaaccacactaga 1758
|||||
Db 4636 CTGCCCAACAAAGAGGTAGAAATGTCCACACAGATACATGCTTACCCACTTAGA 4695
OY 1759 actaaacactgagcagctgtgtcagaacatcccttcaatgttctgaggtaccaga 1818
|||||
Db 4696 ACTAACACCTGACACCACTGTTGTCAGAACATCCCTTATCATGTGTGAGGCTACACGA 4755
OY 1819 gtggtgactcggtgtgtaaacattgtgtgcccgcagcagcgtgttaatgctaactcc 1878
|||||
Db 4756 GTGGCTGACTGGTGTAAACATTGTGTGCTGCAAGCTGTTAATGCTAATTCCTCC 4815
OY 1879 agaatacctcagaagaagagactaagggagacacccagcgctcactgtggaagtgtac 1938
|||||
Db 4816 AGAATACCTCCAGGAAGAGACTAAGGGAAGCCACCGAGCGCTCATGTGGGAAGTGGAC 4875
OY 1939 ttcaactgagctaaagccggctaaatcagggaaacaatatctatgttcttgtagacac 1998
|||||
Db 4876 TTCACTGAGGTAAAGCCGGCTAAATACGGAACAAATATCTATTGTTTGTAGACACC 4935
OY 1999 ttctcagagatgtgtagagcttactcactactaagaagaagattcaaccgtgtgtgctcaag 2058
|||||
Db 4936 TTTCAGGATGGTAGAGCTTATCTACTTAAGAAAGAGACTTCAACCGTGGTGGCTAAG 4995
OY 2059 aaatactgaggaatcttccaagatttgaaatacctaaggaatagaggtcagaat 2118
|||||

Db 4996 AAAATACTGAGAAATTTTTCCAAAGATTGGAAATACCTAAGTAATAGGTCAGACAT 5055
OY 2119 gctcacgttctgcttgcacaaagtaagtcagagacttgcgaagattgggagattgtg 2178
|||||
Db 5056 GGTCCAGCTTTCGTTGCCAGGTAGTCAAGGACTGGCCCAAGATATTGGGATGTGG 5115
OY 2179 aaactgcatltgcatcaagaccccaagctcaggaagaagtagagagaaatgaatgaacc 2238
|||||
Db 5116 AAATGCACTTGTGATACAGACCCCAAGAGCTCAGAGACAGGTAGAGAGATGAATAGAAC 5175
OY 2239 ataaagaagaccccttaccaaattgacacaaagagactgtcaatgatgtatgtgtctc 2298
|||||
Db 5176 ATTAAAGACCCCTTACCAATTTGACACAGAGACTGCGATTAATGATTGATGGCTCTC 5335
OY 2299 ctgccccttgtgtctttaaaggtagaagaaccccttgaagaagtttgagctgacccctat 2358
|||||
Db 5236 CTGCCCTTGTGCTTTTAAAGGTAGGAACACCCCTTGACAGTGTGGGCTGACCCCTAT 5395
OY 2359 gaattgtctacaggggagaccccccgttggcagaanaatgcttgcataatgtctgatal 2418
|||||
Db 5296 GAATTCCTTACGGGGGAGACCCCGCTTGCGAGAAATTCCTTGCACATAGTGTGAT 5355
OY 2419 gtgtgtcttccagaccttgtctctcagctcagagcgtcgtgagtgtgtgagcgaga 2478
|||||
Db 5356 GTGCTGCTTCCAGCCTTGTGCTTGAAGCTCAAGCGCTCGAGTGGGAGGACGCA 5415
OY 2479 gctgtgagacagctcccgaggagcctactcagagagagacttgaagttccacatgctc 2538
|||||
Db 5416 GCGTGAAGACAGCTCCGGAGGCGCTACTCGAGAGAGACTTGAAGTTCCACATCCCTTC 5475
OY 2539 caagttgagatctcagctatgttagacgcacacgtgtgcaagaaacctcagagactcgtg 2598
|||||
Db 5476 CAAGTTGGAATTCAGCTATGTTAGACGCGCACGCTGAGAGAACTCCGAGACTCGTGG 5535
OY 2599 aagggacctatctgtacttcttgacacaaacaaagcgtgtgaaatcgaagaaatcccc 2658
|||||
Db 5536 AAGGACCTTATCTGCTGACTTGTGACCAACCAACGCGTGTGAATTCGAAGAAATCCCC 5595
OY 2659 acctgcatcatcatccacgctttaaagcgcgcacccctccgacttcgggttggaaagcc 2718
|||||
Db 5596 ACCGGAATCCATCATCCACGTTAAAGTGTGGGCCACCTCCGACCTGGGGTGGAGAGCC 5655
OY 2719 gaagaagactgaaataatcccttaaagcttgcctccatcagcgtgtgttcttacttctgaat 2778
|||||
Db 5656 GAAAGACTGAGATCCCTTAAGCTTCCGCTCCATCGCTGCTGCTTACTTACATCAAT 5715
OY 2779 aactcccaagttaatgtgtaaacgcttgtgtgagcagccgaactccataaaccctatc 2838
|||||
Db 5716 AACTCCCGACGCGAGTAGTAACGCTTATAGACAGCTGGAACCCCATAGACCTTTATC 5775
OY 2839 tctcaactgttacttacttactcgttccgtaacaggtatataatataacagcactcaaggga 2898
|||||
Db 5776 CTTACCTGTGCTGATTATTGACCTGATAGCGGTGACACGTAAATAGACATCGAGGT 5835
OY 2899 ggtctcccttggggagcctgtgtgcctgaatataatgtctgtccttgatcaagtaaccctg 2958
|||||
Db 5836 TGCTCTTAAGAGGACCTGTGGGCTGACCTGCTTCTGCTCCGATTAATTAACCCCGC 5895
OY 2959 tctcaatgaccagagccaccccccagatgtactcgtgtgtcaggtttaaagttgtccc 3018
|||||
Db 5896 TGTTA-----AAGACACACTCCCACTAGTCCGTAGTTATGGGTTCTATTGCTGCC 5949
OY 3019 aggaccccaataatgaagaatlatgttgaaatccctcaggaatttcccttgcagaacatg 3078
|||||
Db 5950 AGS---CACAGAGAAGAGAAATACTGTGGGGTTCTGGGAATCCTTGTGAGAGAG 6006
OY 3079 gagctgcataacttcaatgagtaggaaatgtgaaatgtgcagctcctcagcaagacagag 3138
|||||
Db 6007 GAGCTCGTACCTCCAAAGATGAGACTGGAATAGGCGATCTCTCCAGAGACGGGT 6066
OY 3139 aagttacttcttgttaa 3156
|||||
Db 6067 AAAATTCTCTTTGTCAA 6084

Dh 3810 ATAACTGTAAATAGCCCCCATGATGAGAAACATGTTGCGAGCCCCCAGACCGATNG 3869
Qy 919 atgaccaagccccgacatgacacatacaagccgtctctcaagagagggctaccttc 978
Dh 3870 ATATACCAAGCCCGCATGACCAATATCAAAAGCTGCTCTCACAGAGGGGTCACTTTC 3929
Qy 979 gctccacacagccgtctcaacccctgcaactctctgctcgtgaagaaactaagaacagt 1038
Dh 3930 GGTCCACACAGCCGCTCTCAACCCCTGCACTCTTCTGCCGAAGAGACATGATACCACTG 3989
Qy 1039 actcaatgattgcataactatattgattgagagactgggtctcgcaagagaccttaacag 1098
Dh 3990 ACTCAAGATTGGCATCACTATTTGATTGAGGAGACTGGGGTCCGCAAGACCTTACAGAC 4049
Qy 1099 ataccgtactcgtgagaaatgctaaactggttactgactgacgagaaagactatggtgtaa 1158
Dh 4050 ATACCCCTACTGAGAGAAAGTCTTAACCTGTTCTACTGACGGAAGAGCTATGTTGTGAA 4109
Qy 1159 ggtaaagagatgctggtggtggtgagtggtgagacggcagacatctggtggtgagcagc 1218
Dh 4110 GGTAAAGAGATGCTGGGGCGGCAGTGTGAGCGGGACCCGACATCTGGGCCAGCAC 4169
Qy 1219 ctgcccgaaggaactcagcgcaaaagctgagctatggtccctcaacgacttctgctg 1278
Dh 4170 CTGCGGGAAGCACTTCAAGCGCAAAAGCTGAGCTCATGGCCCTCACGCAAGCTTTCGG 4229
Qy 1279 ctgcccgaaggaatctcaataaattatccagacagatagatgctcttggcagctga 1338
Dh 4230 CTGCGCGGAAGGAATTCATTAACATTTATACGACAGAGATGAGCTTTCGCGCTGCA 4289
Qy 1339 caagctacacggggccatcatalaataaagaagggtgtgtactactcagcagagagaaata 1398
Dh 4290 CACGTACACGGGGCATCTATTAACAAAGGGGTTGCTTACTCTCAGCGAGGAGGGAATA 4349
Qy 1399 aagaacaaagaggaattcttaagctatlaaagaccttaactttgccaagaagctagct 1458
Dh 4350 AAGAACAAGAGGAATTTCTAAGCTTATTAGAGGCTTACATTTTGCCAAAAGGCTAGCT 4409
Qy 1459 attataactgtccctgacacatcaagaagcaaatcatalatagaaggagaaagctatg 1518
Dh 4410 ATTATACACTGCTCGAGCATCAGAAAGCCAAAGTCTCATATCTAGAGGGGAACCAAGATG 4469
Qy 1519 gctgacgggtgtgccaagcaagcgccaggtgttaactctctgctataatagaagc 1578
Dh 4470 GCTGACCGGTTGCCAAGCAGGAGCCAGGCTGTTAACTTTCGCCATATATAGAAACG 4529
Qy 1579 cccaagccccgaagaccagacagacatcaacctagaagaactgtgcaagagataaaag 1638
Dh 4530 CCGCAAGCCCGAAGCCAGACGACAGTACACCTTAGAAGACTGGCAAGATTAATAAAG 4589
Qy 1639 atagaccagtctctggaactcggagggagggcgtctataacctatattggaagaaatc 1698
Dh 4590 ATAGACCAAGTTCTGTGAGACTCCGAGGGGACCTGCTATACCTCATATGGAAGAAATC 4649
Qy 1699 ctgcccacaaagaagggttagaatatgtccaacagatacatgtctcaaccacactga 1758
Dh 4650 CTGCCCAACAAAGAGGGTTAGATATGTCCAAGATAGTATGTTTAAACCCACCTTAGGA 4709
Qy 1759 actaaacactgcaagcagttgtgtcaagaacatccctatcatgttctgaagctacaaga 1818
Dh 4710 ACTAAACACCTGACAGCTGTTGTCAGAAACATCCCTTATCATGTGTCTGAAGGCTAACGGA 4769
Qy 1819 gtggtctgactcgtgtgtcaaacattgtgtgctcctgcaagctgttaagtactccctcc 1878
Dh 4770 GTGCTTACTGTGGTTCATAACATTTGTCTCCCTGCGCAGGTGTTATGTATATCTTCC 4829
Qy 1879 agaatactccagaaagagactlaagggagagccaccacagcgctcaactggaagtgagc 1938
Dh 4830 AGAATACTCTCAGGAAGAACTAAGAGGGAAGCCACCCAGGCGCTACTGAGGAAGTGAGC 4889
Qy 1939 ttcaactgaggtlaaagccggttaataatcgaagaacatlatctatgttttggtagaaccc 1998
Dh 4890 TTCACTGAGGTAAAGCGGCTTAATATCGGAACAAATATCTATTTGTTTGTAGACACC 4949

Qy 1999 ttccaagatgggttagaggtcttaccactaagaagaagacttcaacgtgtgtgctaag 2058
Dh 4950 TTTTCAGGATGGGTAGAGGCTTTATCTACTTAAGAAAGAGACTTCAACCGTGGGTGAAG 5009
Qy 2059 aaatactcggaggaatttttccaagatttggaaatcactaagtgaatgaaggtcaagaat 2118
Dh 5010 AAAATACTGGAGGAATTTTTCAGATTTTGAATTTGAAATTTACTTAAGTAATTAAGGTGAGACAT 5069
Qy 2119 ggtccagcttctgttgcacagtglaagcaagtgacaggaactggaagattatgggattctg 2178
Dh 5070 GGTCCAGCTTTCGTTGGCCAGTAAAGTCAAGGAGATGAGGCAAGATTTGGGATTTGATTGG 5129
Qy 2179 aaactgcatgtgcatacagaaccccaagctcaggaaggttagagagatgaatgaagacc 2238
Dh 5130 AAATGCAATTTGATACAGACCCCAAGCTCAGAGCAGGTAGAGAGATGATAGAAC 5189
Qy 2239 attaaagagcccttaccaaattgacacacagagactggaacttaagtattgattgctctc 2298
Dh 5190 ATTAAAGAGACCTTACCAATTTGACCAAGAGACTGGCATTTATGATTTGATGGCTCTC 5249
Qy 2299 ctgccccttgtcttctttaggtgagagacacccctggaagcttgggtcgtgacccctat 2358
Dh 5250 CTGCCCTTGTGCTTTTATAGGTGAGAGAACACCCCTGAGAGTTGGCTGACCCCTAT 5309
Qy 2359 gaattgctcaaggggagaccccccggttggcagaatgtgcttgcacatagtgctgat 2418
Dh 5310 GAATGCTCTACGGGGAGCCCCCGGTTGGCAGAAATTCCTTTGCATATGCTGTAT 5369
Qy 2419 gtgctgtcttcccaagccttgtctcttaagctaaagcgctgagtggtgtagagacga 2478
Dh 5370 GTGCTGTTTCCAGACCTTGTGTTCTAGGCTTAAGCGCTCAGTGGGTGAGCGACA 5429
Qy 2479 gcttggagacagctccggagagcctaactcagggagagacttgaagtccaactgcttc 2538
Dh 5430 GCGTGAAGAGAGCTCCGGGAGGCTTACTCAGAGAGAGACTGCAATTCCACATCCCTTC 5489
Qy 2539 caagtttggagattcagctatgtttagagcgacagctgtcaggaacctcgaagctcgtg 2598
Dh 5490 CAAGTTGGAGATTCACTATTTAGTAAAGCCACCGCTGCGAGGAACCTCGAGACTCGTGG 5549
Qy 2599 aagggaacctatctgcttacttggaccacaaagcgctgtgaagtgaaagaaatcccc 2658
Dh 5550 AAGGACCTTATTTCTGTTACTTTTGGACCAACACGCGGTGAATACTCAAGAAATCCCC 5609
Qy 2659 acctgatalcatgcatcccaagcttlaagcgcgcaaccccgatctgggtgtgaaagcc 2718
Dh 5610 ACCTGATCATGATCCACAGCTTAAGCGCGGCCACCTCCGACTCGGGGTGGAGAGCC 5669
Qy 2719 gaaaagacttgaataatcccttaagcttgcctccatcgctgtgttcccttactgttcaat 2778
Dh 5670 GAAAGACTGAGATCCCTTAAAGCTTCGCTCCATCGGCTGTTCTTACTTAACAT 5729
Qy 2779 aactctcaagttaagtgaagacgcttgtgagacgcccgaactcccataaacctatc 2838
Dh 5730 AACTCCCGAGGCGAGTAAGTAACGGCTTTATTAACAGCTCGAAGCCCAATAGACTTTATTC 5789
Qy 2839 tctcaactggttacttaactgactccggtlaaggtatlaattaaacaacactcaagggga 2898
Dh 5790 CCTTACTGCTGATTTATTTAGACCTTGATAGGGGTGCTACTGTAATTAACACTCGAGGT 5849
Qy 2899 ggtcccttgggagacgtgtgtgctgaattatattgtcgtccttgaagtaaatcccg 2958
Dh 5850 TGTCTCTAAGGACACTGTGTGCTGAACTGATTTTCGCTCGATGATTAAACCCGC 5909
Qy 2959 tctcaatgacagcgcaaccccccgatgactccgtgttcaaggttlaagtttgc 3018
Dh 5910 TGTTA-----AAACACACCTCCCAACCTTCGATGTTATGGTTCTATTGCTGCC 5963
Qy 3019 aggacccccaaataatgaagaatatgttggaaatccctcaagatttcccttgaagcaatg 3078
Dh 5964 AGC-----CACAGGAAGGAATTAATGAGGCGTTCTGGGAATCCCTTGTAGAGATG 6020

QY 3079 gagctgcataactcttaataatggaattggaatgagcagctctcagcaagacaggt 3138
|||||
Db 6021 GAGCTCGTCACCTCCACGATGAGACTGGAATGCGCATCTCTCCACAGCCGGGT 6080
QY 3139 aagttactctttgttaa 3156
|||
Db 6081 AAAATTCTCTTGTCAA 6098

RESULT 9
ARI30475
LOCUS ARI30475 8132 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 3 from patent US 6190861.
ACCESSION ARI30475
VERSION ARI30475.1 GI:14118800
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 8132)
AUTHORS Fishman,J. A.
TITLE Molecular sequences of swine retroviruses method of using
JOURNAL Patent: US 6190861-A 3 20-FEB-2001;
FEATURES
Location/Qualifiers
1..8132
BASE COUNT 2248 a 1977 c 2037 g 1870 t
ORIGIN

Query Match 87.8%; Score 2915.2; DB 6; Length 8132;
Best Local Similarity 94.3%; Pred. No. 0;
Matches 3051; Conservative 0; Mismatches 173; Indels 12; Gaps 2;

QY 22 gatgctcttcttgctctgagatataccccaactaagccaactcttctgcttgaatg 81
|||||
Db 2970 GATGCTTCTTCTGCTGAGATTACACCCACTACCAACCTTTTGCCTTCGAATGG 3029

QY 82 agagatcaggtacggaagaacccggcagctcagctgagaccgagctgccccaaaggttc 141
|||||
Db 3030 AGAGATCCAGGTACCGGAAGAACCGGCGAGCTCAGCTGACCCGAGCTGCCCAAGGGTTC 3089

QY 142 aagaactcccgacacatctttagaagaagccctacacagggacttggccaacttcagatc 201
|||||
Db 3090 AAGAATCCCGACCATCTTTGACGAAGCCCTACACAGAGACCTGCGCAACTTCAGATTC 3149

QY 202 caaacacctcagtgacctctctcagtaagctgagatgacctgctctggcggaagcaac 261
|||||
Db 3150 CAACACCTTCAGGTGACCTCTCTCCAGTACGTGATGACCTGCTGCGGAGCCACAC 3209

QY 262 aacacagactgcttagaagaagctaggaagcactgactgctggaattgtctgacctaggtac 321
|||||
Db 3210 AAACAGAGATGCTTAGAAGGACGAGAGCTACTGCTGGAATTTGCTGACCTAGGCTAC 3269

QY 322 aagacctctgtaagaagcccaagcttgcagagaagagtaacatacttgggtacagt 381
|||||
Db 3270 AGAGCTCTGCTAAGAAAGCCAGATTGTCAGAGAGAGAGTAACTACTGGGGTACAGT 3329

QY 382 ttgcgggggagcagcagatgctgacggaagcaggaagaacacgtgagctcaggtacag 441
|||||
Db 3330 TTGCGGGAGCGGAGGAGTGTGACGGAGGACGAAAGAACTGTGTCAGATACCG 3389

QY 442 gccccaacacacagccaacaaagttagagagcttcttgggagacagctggaatttgcagactg 501
|||||
Db 3390 GCCCCACACACGCAAAATAATGAGAGGTTTGGGGACAGCTGGAATTTGCGAGACTG 3449

QY 502 tggattccggggttgcagacctagcagcccaactctaccgctaaaccaagaagaagag 561
|||||
Db 3450 TGGATCCCGGGCTTGGCGACTTACAGCCCACTCTACCCGCTAACCAAGAAAGAAAGG 3509

QY 562 ggaattctcgtgacctcgtgagaccaagaagagatctgactataaagaagcctgctg 621
|||||
Db 3510 GAATTCCTCTGGGCTCTGAGCACCAAGAGGCAATTTGATGATCAAAAAAGGCGCTGCTG 3569

QY 622 agcgacactgctctggccctccctcagcgttaactaaaccttaccctttagtggatgag 681
|||||
Db 3570 AGCGACCTGCTCTGGCCCTCCCTGACGTACTTAACCTTTATGCTGATGAG 3629

QY 682 cgtaaaggagtagcccgaggaagtttaaccbaacctagaacatgagaggaactgct 741
|||||
Db 3630 CGTAAGAGAGTAGCCCGGGAGTGTAAACCAACCTAGACCATGAGAGACCTGCTG 3689

QY 742 gcttaacctgaagaagctgacctgtagcaggtggttggccggaatctgaagct 801
|||||
Db 3690 GCTTACCTGTAAAGAGCTGATCTGTACCGAGTGGGCGCATATGCTTAAGGCT 3749

QY 802 atcgcaagctggtgcataactgctcaagaagcttacaatgtaacttgggacaacata 861
|||||
Db 3750 ATCGCAGCTGTGGCCATAGTGTGTAAGAGCCCTACCAATTTGATTTGGGACAAATATA 3809

QY 862 actgtaalagccccccatgcatlgagaacatcgttcggcaagcccccaagacgaatgag 921
|||||
Db 3810 ACTGTAATAGCCCCCATGATGATGAGAACATCTGTTGGCAGCCCCAGACCGATGATG 3869

QY 922 accaagcccgcatgaccacatcaagaagcctgtctcagaagaaggtcacttgcct 981
|||||
Db 3870 ACCAAGCCCGCATGACCCACTATCAAGCCTGCTTCAACAGAGAGGTCACGCTGCT 3929

QY 982 ccacacagccgtctcacaacctgccaactctctgcttgaagaagatgataaccagtaact 1041
|||||
Db 3930 CCACCAAGCCGCTCTCAACCCCTGCGACCTTCTGCTGTAAGAAGATGATGAAACAGTACT 3989

QY 1042 catgattgcaatcaactatgattgagagaagctgggttcgcaagaaccttaacagata 1101
|||||
Db 3990 CATGATTGCCATCAATTGATTGAGAGAGACTGGGCTCCGCAAGCACTTACAGACATA 4049

QY 1102 ccgctgactggaagaagtgtaacctggttaacttgaacgaagcagctatgtggtggaagt 1161
|||||
Db 4050 CCGCTGACTGGAAGAAGTCTAAGCTGCTGCTACAGCAAGACACTATGTTGGGGAAGT 4109

QY 1162 aagagatgctctgggagcagtagtgagacgggcccagacatctggcgaagcagctg 1221
|||||
Db 4110 AAGAGGATGCTGGGGGGGGGCTGTGTGAGCGGAGCCGCAAGATCTGGCGACAGCTG 4169

QY 1222 ccggaaggaacttcaagcgaagaagctgagctatgagccctacgcaagcttgcgctg 1281
|||||
Db 4170 CCGGAAGGAATTCAGCACAAAAGCTGAGCTCATGCGCTTCAGCGAAGCTTGGCGCTG 4229

QY 1282 ggcgaaggaagaaatccataaacttatacgaagcagagatgacttctgagactgcaac 1341
|||||
Db 4230 GCCGAAGGGAATTCATTAACATTTATACGACAGCAGGTATGCTTTGGACTGCACAC 4289

QY 1342 gtacacggggccalcataaacaagaagggttgcctaccccaagcaggaaggaataaag 1401
|||||
Db 4290 GTACATGGGGCATCTTAACCAAAAGGGGTGCTTACCTCAGAGGAGGAGGAATTAAG 4349

QY 1402 aacaagaaggaatcttaagcctatagaagccttaccattggccaagaagctagctatt 1461
|||||
Db 4350 AACAAAGAGGAATTTCTAAGCCTATTAGAGCCCTTACTAATTACCAAAAAGGCTAGCTATT 4409

QY 1462 ataacgttctgacatcgaagaagcgaagaatctatcttagaggaagcagaatgct 1521
|||||
Db 4410 ATTAAGTCTCTGACATTCGAAGAGCTTAAGATCTCATATCAAGAGAAACCGAGATGCT 4469

QY 1522 gaccgggttgcgaagcagcagccagctgttlaaccttctgctataatagaaacccc 1581
|||||
Db 4470 GACGGGGTGTGCAAGGAGGAGGCCCAAGGTTTAACTTCTGCTTAATATAGAAATGCC 4529

QY 1582 aaagccccaagaaccagaacgaaglacacccctagaagaactggaagaagataaaagaata 1641
|||||
Db 4530 AAAGCCCAAGAACCAACAGACATACACCTTGAAACATGCGAAGAGATAAAAAGATA 4589

QY 1642 gaccagttcttgaagctcggagggagactgctatcccatatgggaaggaatccctg 1701
|||||
Db 4590 GACAGTTCTCTGAGACTCCGGAAGGAGCCTGCTATACCTCAAGTGGGAAGAAATCTCTG 4649

QY 1702 ccccaagaagaaggttgaatatagtccaaagatacatcgtctaaacccactaggaact 1761
 |||||
 Db 4650 CCCCAAGAAGAAGGTTGAATATGTCACAGATACATCGTCAACCCACTAGGAAC 4709
 QY 1762 aaacaccctgcaagcttgatgaacaccccttactaattcttgagctaccagagtg 1821
 |||||
 Db 4710 AAACACCTGCACATTTGGTCAGACATCCCTTATCATGTTGAGGCTACCGAGGATG 4769
 QY 1822 gctgaactcgttgatcaaatgtgtgcccctgcagctggttaatgtaactctccaga 1881
 |||||
 Db 4770 GCTGACTCGGTGTCAAACATTTGTGCTGCCAGTGTGTAATGCTAATCTCCAGA 4829
 QY 1882 atactccaagaagaagaactaagggaagcaaccagagctcactgtggaatggaactc 1941
 |||||
 Db 4830 ATGCTCCAGGAGAGAGACTTAAGGGAAGCCACCGAGCGCTCATCTGGAGAGTGACTTC 4889
 QY 1942 actgagtaagaagccgagctaaataagcaacaataatactattggttttggatgaacctt 2001
 |||||
 Db 4890 ACTGAGTAAAGCCGGCTTAATACGGAACAAATACCTATTGTTTGTAGACACCTTT 4949
 QY 2002 tcaagatggttagaaggttatactactaagaagaagacttcaacgctgtgtcgaaga 2061
 |||||
 Db 4950 TCAGATGGGTAGAGGCTTATTCCTACTAAGAAAGAGACTTCAACCGTGTGGTAAAAA 5009
 QY 2062 atactggaggaatttttccaagatttggaataactaagtaataaggatcagaatggt 2121
 |||||
 Db 5010 ATACTGGAGAGAAATTTTCCAAAGATTGGAATACCTAAGTAAATAGGCTCAGCAATAGT 5069
 QY 2122 cgaagcttcgttgcccaagtaagtaaggagacttgcccaagatattggagattgtgaa 2181
 |||||
 Db 5070 CCAGCTTTTGTGTTGCCAGTAAAGTACAGGACTGGCCAAAGATATGGGATTTGATGGAA 5129
 QY 2182 ctgcattgtgatacagaccccaagctcaggaagctaggaagatgaatagaacatt 2241
 |||||
 Db 5130 CTGCATGTGCATACAGACCCCAAGCTCAGACAGGATAGAGAGATGAATAGAACATT 5189
 QY 2242 aaagagacaccttaccacaaattgacacagagactgagctaaatgattgagtgctcc 2301
 |||||
 Db 5190 AAAGGAGACCTTACTAAATTTGACCGGAGACTGGCCCTTAATGATGATAGCTCTCTG 5249
 QY 2302 ccccttgctgcttttaaggtagaagaaccccttgagacagtttgagcgagccctatgaa 2361
 |||||
 Db 5250 CCGTTTGTGCTTTTAAAGGTTAAGAAACACCCCTGACAGTTTGAGGCTGACCCCTATGAA 5309
 QY 2362 ttgctctacggggagaccccccgcttgagcagaatattgctcttgacatagtgtcgtatg 2421
 |||||
 Db 5310 TTACTCTACGGGGAGACCCCGCCCATTTGGTAGAAATTTGCTGTACATGATGCTGAGCTG 5369
 QY 2422 ctgctttcccaagcttctgtctcctagctcgaagcgctcgagtggtgagcagagagcg 2481
 |||||
 Db 5370 CTGCTTTGCCAGCCCTTGTGTTCTTACAGCTCAAGGCACCTTGAAGTGAGTGGAGCAACGAGCG 5429
 QY 2482 tggagagacgctccgggagagctactc--aagagagagacttgcaagtccaactcgcttc 2538
 |||||
 Db 5430 TGGAGGCACTCCGGGAGGCTACTCAGAGAGAGAGAACTTGGAGATCCACATCGTTTC 5489
 QY 2539 caagtgtgagaattcaagctatgtttagaagcagcagcgctgaggaacccctcgagactcg 2598
 |||||
 Db 5490 CAAGTGGAGATCAGTCTACTAGCTTAGAGACGACCGTGCAGAAACCTCGAGACTCGTGG 5549
 QY 2599 aagagacttatctcgtacttttgacacacacacagcgctggaagtgcgaagaatccccc 2658
 |||||
 Db 5550 AAGGAGCCCTTAATCTGACTTTTGAACACCAACAGGCTGTGAAGAGTGAAGATCTCC 5609
 QY 2659 accgtgattcaatgcatcccaagctaaagcagcgccactccgattcgagggttgaagacc 2718
 |||||
 Db 5610 ACCTGGATCCATGCTCCACGTTAAACGCGGACCACTCCGATTTGGGGTGGAAAGCC 5669
 QY 2719 gaaaagactgaaaaatcccttaagcttcgctccatcgctgtgttccctactcgtcaat 2778
 |||||
 Db 5670 GAAAGAGACTGAAATCCCTTAAGCTTCGCTCCATCGCTGCTTACTCTGCTCAAT 5729
 QY 2779 aactccccaagttaatgtaaaagccttgtagacagccgaactccataaaccttatac 2838

Db 5730 AACCTCTACAGCTAATATGATGCGCATAGAGAGACACCTGAATCCCATAAACCTTATC 5789
 QY 2839 tctcaactgttacttaactactcgcggttaaggtatataataaagaactaaaggga 2898
 |||||
 Db 5790 TCTCACTGTTATTAATCTACTCTCGGACAGGATTAATATCAACAAACCTCAAGGGA 5849
 QY 2899 gctcccttgaggagacctggtgacctaatatatactgctcctcgaatcaatccctgg 2958
 |||||
 Db 5850 GCGTCCTTAGGAACCTGGTGGCTATCATACGTTTGGCTCAGATCAATTTCTCTAG 5909
 QY 2959 tctcaatgaccagggccacaccccccgatgactcgtgtcttgaagtttaagttgccc 3018
 |||||
 Db 5910 TCT-----GACCTCACCCCGAGATATCTCCATGCTCAACGAGATTTATGTTGCC 5960
 QY 3019 aggaaccccaataatgaagaatatgttggaatccctcaggtttcttcttgcaagagatg 3078
 |||||
 Db 5961 AGGACCACTAATTAATGGAATATGCGGAATATCCAGAGATTTCTTTGTAACCAATG 6020
 QY 3079 gaactgataactctcaatgatgtggaattggaatgagccagctcctcgaagaagagat 3138
 |||||
 Db 6021 GAACGTGTAACTCTAATGATGATTTGAATGCGCACCTCTCACAGAGATAGGT 6080
 QY 3139 aagttactcttgttaacaaatccatccagttataatcaatlaattatgycatggag 3198
 |||||
 Db 6081 AAGTTTTCTTATGTAACACCTATACAGCTCTCGACATTAATTAATTAACCTGAGAT 6140
 QY 3199 atggaagattggaacaggggttacaagaagtgtgagaatlaagcaataagct 3254
 |||||
 Db 6141 TAGAATCGAAGCCCAAGTGTCTCTTACACACTAGATTAATTAATTAATGTT 6196

RESULT 10
 LOCUS AF038600
 DEFINITION Sus scrofa porcine endogenous retrovirus PERV-MSL mRNA, complete sequence.
 ACCESSION AF038600
 VERSION AF038600.1 GI:3133301
 KEYWORDS
 SOURCE
 ORGANISM
 pig.
 Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 1 (bases 1 to 8132)
 Akiyoshi,D.E., Denaro,M., Zhu,H., Greenstein,J.L., Banerjee,P. and Fishman,J.A.
 TITLE Identification of a full-length cDNA for an endogenous retrovirus of miniature swine
 JOURNAL J. Virol. 72 (5), 4503-4507 (1998)
 MEDLINE 98216827
 REFERENCE 2 (bases 1 to 8132)
 AUTHORS Fishman,J.A.
 TITLE Direct Submision
 JOURNAL Submitted (16-DEC-1997) Infectious Disease Unit, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02114, USA
 3 (bases 1 to 8132)
 Fishman,J.A.
 TITLE Direct Submision
 JOURNAL Submitted (15-MAY-1998) Infectious Disease Unit, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02114, USA
 REMARK Sequence update by submitter
 COMMENT On May 16, 1998 this sequence version replaced gi:3116441.
 FEATURES
 source
 1..8132
 /organism="Sus scrofa"
 /strain="miniature swine"
 /db_xref="taxon:9823"
 /cell_type="lymphocytes"
 /note="porcine endogenous retrovirus PERV-MSL"
 LTR 1..564
 CDS 585..2159
 /codon_start=1

/product="gag protein"
/protein_id="AAC16763.1"
/db_xref="GI:3116442"
/translation="MGQIVTTPSLTLDHMTVEKSAHNLISVQYKKGPMQTCVSEMP
TEPVQWSESTSEITLAKAVITOTGSHSDDEPLYLTIMODLDEPPVWPKWLN
KPKPGRIALIEENKSHAEKVPSPHLYPELEPPAPMEPSVPPYLAQAGRG
FPAPGAPAVEPFAAGTGRKRGATPERTDEIATLPRTGPPPGQOLPQLYWPS
ADLYNMKTNHPESDEPRLTGIVESLMEFSDPTWDCOOLLOTLETERERILIEA
RKVPGADGPRITLONEIDMGFLPLRGMIDYNTAEERESLKITRYALVALITGSADIR
TNLAKEVWOGNPPPSVLEPLLEAPRYTPEDPSYKROKASERVALEITGSADIR
KTIQRLGLOEALRDLVKAKEYVYIKRETEERERKREBERERENKROEKULI
KILAAVEGKSNKEDRDFKTRSGRSGNIGNRTPLDKDOCAVCERGHMARNCPK
KGNKGRILALEEDND"
2307..5744
/codon_start=1
/product="pol protein"
/protein_id="AAC16764.1"
/db_xref="GI:3133302"
/translation="MGATGQOQYPTWTRRTVLDGVRVTHSLVIEPCFAPLLGRDL
TKMAQISEQCKPEVSANNKPTVTLQDDEYRLYSPLVDPDNIQWLEPQPM
AETAGMGLAKOVPVOYLKASATPVSVROYPLSKAGSIEPIPHQRLIOGILVPO
SPMNTPLPYRKPSTNDYRVDLREVNKRVODIHTVYNPNTILCALIPROSKWTL
DLKDAFECRLHTPSQPLAFEMRPDPTGRTQOLITPLPOGFKNSPTIEDALRDL
ANFRIOPOVTLQIVDLDLAGATKODCEKALILESDLYRASAKKAOICRE
VYTLGSLBDGQWLTBAKKTVOQIPAPITAKOMERLIGTAFQCLMIPGATLAAP
LYPLTEKEGEFSWAPRHOAFPAIKKALLSAPALAPDVTGKPTLYDREKIVARGVL
TOTLGPWRPVALYSKLDLPVAGMPICLKALAAVALIVKADKLITGONITVAPHA
LENTVROPDRMNTNARMTHYOSLITLERTFAPPAALNAPLLEPEDEPVTNCHQ
LLEETGVRKDLDTPLTGEVLTWPTDSSYYVEGRMAGAAVVDGRTIMASSLPQG
TSAQKELMALTOALRLAEGKSINITYDSVAFATVHGALEYKQGLTTSAGRIK
KEBILSLAEVHLPRKALIHCPGQKADLISRGOMADRYAKQAQOVNLPTTEM
PKAPERRQVLTLEMOEIKRIDOFSETPEGTCTSDGKELILPKHGLEVOQIHRLLH
LGTKLHLOQVARTSPYHRLPGVADSVHGCYCOLVANNPSPGKRLRSHGHA
MEYDFTEVPKAGNKYLLFVDTESGWEAVPTKSTSTVAKKILIEIPRERIPK
VIGSDNGPAPVAHVSOGLAKIIGIMKLCIAYRPOSSGVEEMNRTIKETLIKLAET
GVNDWIALPLPULFRYRANTPGQGLTPLYLLAGPPLVLEIASADLLSOPLESR
LKALEWROKRAMSOLKEATSGGDDLIPIRFQGVSVYRRHRAENGLRWMGPLYLT
LTPPTAVKVGISTWIMASHVKKPAPPDGMAKENTENLKLRLHRYVYVSNNLSD"
5620..7536
/codon_start=1
/product="env protein"
/protein_id="AAC16765.1"
/db_xref="GI:3133303"
/translation="MHPILNRRLPLRGKPRRLKIPLSFASIAFWLTLSTQNGM
RIEDSLNSHKPLSLTLITDSCGININNOGEAPITGMPDYLCISVLSLSP
DILHAGFYVPCPPNKGKHCNPDEFCOMKONCTSNMGVYKPTSOODRSESVN
TYTSGOFNLTMRIGSPKSPSDDLYKISPTKRGKENTLKVNKMSWAMYVGG
SKQPSILITRIKINOLEPMAIPNTVLTORPTQGFSGSNITSGSDPTESST
TKMGALFSLIOGAFOLNSTPTEATSCWLCIASGPPYEGSMARGFNVKHERDQ
CTWGSQNKLTLELSESGKTCIGKVPSSHQHLNHEAFNOTSESOYLVEGYDWMACN
TGLTPCVSTLIVNOTKPCIMVOIVPRVYYPEKAILDEYDRNHRKREPSTLAV
MLGIVAGVGTGTALVLPPOOLLETGSLNHRITEDOLSNKESVNLSESTLSE
VVLONRGGDLLEFKEGGLCVALKKECCGYVDSHSAIDRSNKLRLERLKKRREKRT
OGMEFEMNRSIMLALTLALSALTGPLVLLLTITVPCITINKLIATIRERISAVQIMVL
RQIYOSPSREKMR"

LTR
BASE COUNT 2247 a 1977 c 2037 g 1871 t
ORIGIN

Query Match 87.7% Score 2912; DB 4; Length 8132;
Best Local Similarity 94.2% Pred. No. 0;
Matches 3049; Conservative 0; Mismatches 175; Indels 12; Gaps 2;

22 gatgcctcttcgctgagatatacccccactagcaaacacttttgccttgcgaatg 81
Db 2970 GATGCCCTTCTTCTGCTGGAATTAACACCCACTAGCCAACTTTTGTGCTTCGATGG 3029
QY 82 agaatccaggtacggaagaacccgagctcaacttgcacccagctgcacaaaggttc 141
Db 3030 AGAAGATCCAGGTACGGGAAGAACCCGGGACAGCTCAGCCAGCTGCCCCAAGGTTTC 3089
QY 142 aagaatccccgaacacattttgaagaagccctacagaaggaacttgcacaaactcagatc 201

Db 3090 AAGAACTCCCGGACCATCTTTGACGAAGCCCTACACAGAGACTGGCCAACTTACGAGTAC 3149
QY 202 caaacccctcgaagtgcacctcctcgaatcgtgtagtaccctgccttcgaggaagccacc 261
Db 3150 CAACACCCCTCAGGTGACCTCTCCAGTACGGGATGACCTCTTGTGGGGAGCCACC 3209
QY 262 aaacagaactgtctaaagaatcgaaggaacactgcctgcgaattctcgaactgaagctac 321
Db 3210 AAMCAGACTGCTTACGAAGGACGAAGGACTACTGCTGGAATTTGCTACCTAGGCTAC 3269
QY 322 agagccctcgtcctaagaagcccaagatttcgaagaagaggttaacatacttgggtgacagt 381
Db 3270 AGAGCCTCTGCTACGAAGGCCCAAGATTTCAGAGAGAGAGACTTAACATCTGGGGTACAGT 3329
QY 382 ttgcggagcgagcagcagatgctgcagcagagcagcaggaagaanaactgtagtccagatccg 441
Db 3330 TTACGGGAGCGGCGACGATGCGTGAACGAGGACGGAAGAAACCTGTAGTCCAGATACCG 3389
QY 442 gccccaaccaacagcaagaatgaagaagttttggggagacagctggaatttgcagactg 501
Db 3390 GCCCCACACACACCAACAAATGACAGAGTTTTGGGACAGCTGATTTTGCAGACTG 3449
QY 502 tggatccgagggtttgcagacttgaagcccaactcaccgctlaaccacaaagaagaagg 561
Db 3450 TGGATCCCGGGGTTTGCAGACTTAGAGCCCACTCTACCCGCTAACCAAGAAAGAGG 3509
QY 562 ggaattcctcgtggtcctcgtgagcaccagaagagcatttgaatgatacaaaaagccctgtg 621
Db 3510 GAATTCCTCTGGGCTCCTGACGACCAAGAGGATTTGATGATCAAAAGGCTCTGCTG 3569
QY 622 agagcaaccgtgcctcgtgacctcctcgtgacgttaacaccccttaaccttatgtgtagag 681
Db 3570 AAGCAGACTGCTGCTGCGCTCCCTGACGTAACCTTAACCTTTATGTGGATGAG 3629
QY 682 cgtlaaggaagtagcccgaggaagttttaaaccaccccttagagcacttgaaggaagcctgtt 741
Db 3630 CGTAAAGGAGTGTAGCCCGGGAGTTTAAACCCAACTTAGACCATGAGAGAGACTGTGC 3689
QY 742 gctctcgtcgaagaagatttgatcctgtgaagcaatgtgttgcggtatgtctgaaggt 801
Db 3690 GCTTACCTGTCAAAAGACTTCATCTCTGTAGCAATGTGTGCTCATGTGCTGAAGGCT 3749
QY 802 atcgaagctgtgcatcacttgcgaagaagctgtgcaaatbtaactttggagacaata 861
Db 3750 ATCGCAGCTGTGCTCATACTGTTCAAGAGAGCTGTGCAAAATTGACTTTGGGACAGATATA 3809
QY 862 actgttaataagcccccaatgcatlttgagaagaacatcgttcgagcagccccagacagatgata 921
Db 3810 ACTGTAAATAGCCCCCATGCAATTGGAAGAACATCGTTGCGGACGCCAGACGATGATG 3869
QY 922 accaagcccggaataaccacataaagaagcctgccttcaagaagaaggttaacttgcct 981
Db 3870 ACCAAGCGCCGATGACCCACTATCAAAAGCCTGCTTCTACAGAGAGAGGATGACGTTCCCT 3929
QY 982 ccaccagccgctcacaacctgcacactcttcgctcgaagaagcagatgaacacagtgact 1041
Db 3930 CCACGACCGGCTCTCAACCTTCCACACTCTTGCCTGAAGAGACTGATGAAACAGTACT 3989
QY 1042 catgattgcatcaactatgattgaagaagacatgggttcgcaaggaacttaccagacata 1101
Db 3990 CATGATTGCAATCAACTATTGATTGAGAGAGACTGGGGTCCGCAAGACCTTAACAGACATA 4049
QY 1102 ccgctgactcgaagaagtcctaaccctggttcaacttgcaggaagcagctatggttggaagt 1161
Db 4050 CCGCTGACTGGAGAGTGTAACTGTGTTCACTGACGAGAACAGCTATGTTGTTGGAAGT 4109
QY 1162 aagagatgagcttggaggcgagtgatgtagcaggaacccgacagatctggcgagcagcctg 1221
Db 4110 AAGAGATGAGCTGGGGGCGGGTGTGTGACGAGGAGCCGACGATCTGGGCCAGCAGCTTG 4169
QY 1222 ccggaaggaacttcagcgcaaaaggctgaagctcatgagccctcgaagcagcttgcagctg 1281

Db 4170 CCGGAGAGAACTTACGACCAAAAGGCTGAGCTCAWGGCCCTCAGCGAAGCTTTGGGCTG 4229
 QY 1282 gccgaaggaatccataaacaattatacgcagcagatgcttgacgtcac 1341
 Db 4230 GCCGAGGGAATTCATTAACATTTATACGACAGAGATATGCTTTGGAGCTGCACAC 4289
 QY 1342 gtacacggggccatctataaacaagggtgtcttaccctcgaaggaggaaataaag 1401
 Db 4290 GTACATGGGGCATCTATMAACAAGGGGGTGGCTTACCTCAGACGAGGGAATTAAG 4349
 QY 1402 aacaagaagaatcttaagccctatlagaaagccctatcatcttgcccaaaagctacgtcatt 1461
 Db 4350 AACCAAGAGAAATTTCTAAGCCTATTAGAAAGCCGTACATTTACCAAAAAGGCTAGCTATT 4409
 QY 1462 atacatgctccttgacatacgaagaacacatcatalctagaaggaaacagatgct 1521
 Db 4410 ATACACTGCTCTGACATCAGAAAGCTAAAGATCTCATATCCAGAGAAACCAAGATGCT 4469
 QY 1522 gaccgggttgcccaagcagcagccaggtgttaaccccttgccataatagaagcgc 1581
 Db 4470 GACCGGGTTGGCAAGCAGGCGCCAGGGGTATTAACCTTGCTTATTAATAGAAATGCCC 4529
 QY 1582 aaagccccaagaccagacagacagatcacccctagaagaagctggcaagagataaagaata 1641
 Db 4530 AAAAGCCCAAGACCCAGACACAGCACTACACCTTAGAAGACTGGCAAGAGATTAAGAAAT 4589
 QY 1642 gaccagttctctgagactccggaaggagactgtcatatccatcatatggaaagaatcctg 1701
 Db 4590 GACCAGTTCTCTGAGACTCCGGAAGGAGACCTGCTATACCTCAGATGGGAAGAAATCCTG 4649
 QY 1702 ccccaagaagaaggttagaataatgtcccaacagatacatgtcttaacccactagaagact 1761
 Db 4650 CCCCAAGAAAGGGTTAGATATGTCCAACAGATACATGCTTAACCCCACTAAGAGAACT 4709
 QY 1762 aaacactgacagctgtgtcaagaacatccctcatcatgtcttgaaagctacccaagagt 1821
 Db 4710 AAAACACTGCAGAGCTGTGTGACAGACATCCCTTATCATTTCTGAGGCTACCAAGAGAGT 4769
 QY 1822 gctgactcgtgtgtcaaacatgtgtgcctgcagctgtgttaagttaactccttcaga 1881
 Db 4770 GCTGACTCGGTGTGCAAACTTGTGTGCTCCCTGCAGCTGGTATATGCTTAATCCTTCAGA 4829
 QY 1882 atactctcagaagaagaacacgaaggagacccagcgctcactgaggagagtgagcttc 1941
 Db 4830 ATGCTCTCAGGGAAGAACTAAGGGAAGCACCAGCCCTCAGCTGGGAAGTGACTTC 4889
 QY 1942 actgaagtaaaagccggtlaaataacgaagaacataltcatctgttcttgtagaacctt 2001
 Db 4890 ACTGAGGTAAAGCCGGCTTAATACGGAAACAATACCTATTGGTTTTGTAGAACCTTT 4949
 QY 2002 tcagatgtggttagagctctactactaagaagaagactcaaccgtgtgtgctaaagaa 2061
 Db 4950 TCAGAGTGGGTAGAGGCTTATCTCTACTAAGAAAGAGACTTCAACCGGTGGCTTAAGAAA 5009
 QY 2062 atactgagaagaatcttcccaagattggaatccctaaagtaataaggttcagaacaatgt 2121
 Db 5010 ATACTGGAAGAAATTTTCCAAAGATTGGAAATACCTTAGGTAATAGGCTAGCAATGCT 5069
 QY 2122 ccagcttgcgtgccaggttaagtcaggagactggccaagatalttgggagatgtatgaa 2181
 Db 5070 CCAGCTTTGTTGGCCAGTAAAGTCAAGGACTGGCCAAAGATTTGGGGATTGATTGAAA 5129
 QY 2182 ctgcatgtgtcatatacagaccccaagagctcagacaggttagaaggatagaaccatt 2241
 Db 5130 CTGATTTGTCATACACCCCAAGCTCAGAGAGGTAGAGGATGAATTAACCATTT 5189
 QY 2242 aaagaagcccttaaccaatctgacacagagactgacatlaaattgagatgagctccgt 2301
 Db 5190 AAAGAGACCTTACTTAATTTGACCGGAGACTGGCTTATATTTGATAGCTTCTCCTG 5249
 QY 2302 cccttgtgctttttagagtgagaaacacccctgagacagtttggcttgagacccctatgaa 2361
 Db 5250 CCTTTGTGCTTTTAGGTTAGAGACACCCCTGAGACAGTTTGGGCTGACCCCTTATGAA 5309

QY 2362 ttgcttaagggggaccccccggttgagagaatltgcttgcacatagtgctgagt 2421
 Db 5310 TTACTCTACGGGGAGACCCCCCATTTGTAAGAAATTTCTTGTATAGTCTCATGTG 5369
 QY 2422 ctgcttcccaagccttgttctcttaagctcaaggcgctcagtgagtgagcagcagcg 2481
 Db 5370 CTGCTTTCCAGCCCTTGTTCCTTAGGCTCAAGGCACTTGAATGGGTGAGCAACGAGCG 5429
 QY 2482 tggaaagcagctccggagagccctacc-aggagagacttgcagttccacatgcctc 2538
 Db 5430 TGGAGGCAACTCCGGAGGCTTACTCAGAGAGAGAGACTTCCAGATCCCATGCTTTC 5489
 QY 2539 caagtgtgagatcagctatgttagacgcacacgttgcaagaacactcagacactcgtg 2598
 Db 5490 CAAGTGGAGATTCAGTCTACGTTAGACGCCACCGTGAGAAACCTCGAGACTCGTGG 5549
 QY 2599 aaggaacattactcgtacttltgacacacacacagcgctgtgaaagtcgaagaatccc 2658
 Db 5550 AAGGCGCTTATCTCTACTTTTGAACACACACACGCGCTGTGAAAGTGAAGATTCCTC 5609
 QY 2659 acctgatacatgcatcccaagcttaagccggcgccacccctccgacttggggtggaagcc 2718
 Db 5610 ACTGATTCATTCATCCACCTTAACCGCGGCCACCTCCGATTCGGGGTGGAAAGCC 5669
 QY 2719 gaaaagactgaaatcccttaagcttgcctcatalcggtgtgttcttacttgcatt 2778
 Db 5670 GAAAACACTGAAATCCCTTAAAGCTTCGCTTCAATCGGCTGATTCCTTACTCTGTCAAT 5729
 QY 2779 aactcctaagttaatgtgaaagcgctgtgtgaagcccgaaactccataaaccttalc 2838
 Db 5730 AACCTTCTGACTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5789
 QY 2839 tctcaactgttacttacttacttacttacttacttacttacttacttacttacttactt 2898
 Db 5790 TCTCAGCTGTGTTAATTACTGACTCCGGCAGAGTTTATATACACACACTCAAGGGA 5849
 QY 2899 ggtcccttgggagactgtgtgctgaattatagltcgtccttgcatacgaatccctg 2958
 Db 5850 GGTCTCTTATGAGAACCTGTGTGCTGATCTATACCTTTGCTGCAATGATGATGATGAT 5909
 QY 2959 tctcaatgacagcgagcagcccgatgtactcgttcttaagggttttcgttggcc 3018
 Db 5910 TCT-----GACCTACCCCGAGATATCTCATGCTCAGGATTTTATGTTGCTC 5960
 QY 3019 aggaaccccaataatgaagaatatttgaagaatccctcagaattccttccaagaatg 3078
 Db 5961 AGCACACCAATTAATGGAATTTGGGAAATCCAGAGATTTCTTTGTAACAATG 6020
 QY 3079 gactgtcataacttctaattgaggaattggaatgtgcccagctctcagaagaagagct 3138
 Db 6021 GAACGTGTATACCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6080
 QY 3139 aagttactcttltgttaaacatccctcaagatlaataatlaattatgccaatgag 3198
 Db 6081 AAGTTTCTTATGTCACACCTATACCAAGCTCTGGAATTTAATTACTGACTGTGAT 6140
 QY 3199 atggaagaattggaacagcggttacaaaagatgtacgaataaagcaataagct 3254
 Db 6141 TAGAACTGTGAACCCCAAGTGTCTCTTACAGACTTATACCTAATAAATAAGTT 6196

RESULT 11
 AF038599 8132 bp mRNA MM 21-MAY-1998
 LOCUS AF038599
 DEFINITION Sus scrofa porcine endogenous retrovirus Tsukuba-1 mRNA, complete
 sequence.
 ACCESSION AF038599
 VERSION AF038599.1 GI:3133300
 KEYWORDS
 SOURCE pig.
 ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

	Best Local Similarity	94.1%	Pred No. 0:	Matches 3047:	Conservative 0:	Mismatches 176:	Indels 15:	Gaps 5:
Y	22	gagcctctctcttgccctgagattacaaccacctaagccaacacttlttgcttgaatg	81					
b	2976	gAGGcCTCTCTTgcGcCTGAGATTACACCCCACTAGGCAACACACTTTTggCTTCCAGATGg	3035					
Y	82	agagatccaagtlacgagaaagaccgagcagcttaccttgaccgcgaactgcccgaaggttc	141					
b	3036	AGGAGTCCAGGTCAGGGAAACCGGGACACTGCACCTGGACCcCGACCGCCCAAGGGTTTc	3095					
Y	142	aagaactccccgcagcacttlttgacgaagaccctacacaaaggacttggccaacttcagatc	201					
b	3096	AAAGACTCCCGGACCATCTTTGACGAAGCCCTTACAGAGAACTGGGCCAACTTCAGATTC	3155					
Y	202	caacacctcaagtlgascctctctccagtcaglygatgacctgtcttccttgccgagagccacc	261					
b	3156	CAACACCCCTCAGCTGACCCCTCCCTCCAGTACGTAGTGAAGACCTGCTTGTGGCGGAGACCC	3215					
Y	262	aacacgagactgcttaagaagtacgaaggacactctctggaattgtcttacctaaagctac	321					
Y	3216	AAACGAGACTGCTTTAAAGGCAAGGAAGGACTACTCTCTGAAATTGCTGCTTACCTTCCCTTC	3275					

442 gcccgaaaccagacgcacaaacgaatgaagaaatttttggaaacactgaattttgcacactc 501
 5530 tttacggagcgggacagcgaatgcctgacccagggcaccggaaagaaacitgtagtccagataccg 339

Db 3396 GCCCACAACCAGCCAAACAATGAGAGATTTTTGGGGACAGCTGGATTTCAGACTTC 345

Accession	Sequence	Position
Db	34-36 TGGATCCCGGGGCTTTGGACCTTGGACGCCCCCATCTACCCCGGTAAACCAAGAAAAAGG	3511
QY	562 ggaattcctcctgagcctcctgagcaccagaagcaattgatgctatcaaaaagggccctgcg	621
Db	3516 GAATTCCTCTGGGCTCCTGAGACACAAAGACATTGATGCTATCAAAAAAGGCCCTGCG	3578
QY	622 agcgcacctgctcgcgcctccctcgaaglaactaaaccccttaaccttatgtgataag	681
Db	3576 AGCGACCTCTGCTGGCCCTCCGACGACTAAACCTTTACCTTTATGTGATGAG	3633

QY 682 cgtaaaggatagcccgagagatttlaacccaacctagacatggaggaacctgt 741
|||||
Db 3636 CGTAAGGAGTAGCCCGGGGGTATTAAACCAACCTAGGACATGGAGAACCTGTC 3695

QY	742	gccactactggtcaagaagacttgatccctgtatgacagatggtgttgccgtagtcttgtaagct	801
Db	3696	gccactactggtcaagaagacttgatccctgtatgacagatggtgttgccgtagtcttgtaagct	3755
QY	802	atcgacagctgtgacacatcgtgtcaagaagacgctgcacaaattacatttgagac-agaatat	860
Db	3756	atcgacagctgtgacacatcgtgtcaagaagacgctgcacaaattacatttgagacaaacaaat	3815
QY	861	aactgtataatgccccccatgcatcttggaacaatcglttggcagccccccagaccgatgat	920
Db	3816	aactgtataatgccccccatcatttggaagaacatcglttggcagccccccagaccgatgat	3875
QY	921	gaacaaagcccgcatgtgacccaactataaagccctgctctcacagagaagtgtaacttgc	980
Db	3876	gaacaaagcccgcatgtgacccaactataaagccctgctctcacagagaagtgtaacttgc	3933

QY	981	tcacacagcgcgtctcaacccctgcacattcttcgcctgaagagactgtagaacccagtgac	1040
Db	3336	TCGACCAACCGCTCAACCCCTCCACTCTTCTGGCTGAAGAGACTGATGAACCAAGTAC	3995
QY	1041	tcagatgtagcatcaactatgtaggaagatgggtccgcaagacctaagacat	1100

Dh 3996 TCATGATTGCCATCACTATTGATGTAGAGAGACTGGGGTCCGCAAGACCTTACAGACAT 4055
Qy 1101 accgcgtgactggaagtgcttaacctggttcaactgaacggaagcactatgltgtgaag 1160
Dh 4056 ACCGCTACAGGGAAGTGTAACTGGTTCACCTGACGGAAGCAAGCATGTGTGGTGAAG 4115
Qy 1161 taagagatgctggggcgagctgggtgaacgagaccgacatctggccgagcct 1220
Dh 4116 TAAGAGATGCTGGGGCGGGGTGGTGAACCCGACATCTGGGCCAGCGCCT 4175
Qy 1221 gccggaaggaactcagcgcaaaagctgactatgcccctcaagcaagcttgcgct 1280
Dh 4176 GCCGGGGAAGCTTTCAGCAAAAAGCTGAGCTCATGGCCCTCACCAAGCTTTGGGCT 4235
Qy 1281 gccggaaggaatccataacattatagacagcagatgactgcttggcactgaca 1340
Dh 4236 GGGCGAAGGGAATTCATTAACATTTATACGACAGAGATAGCTTTGGCACTGCACA 4295
Qy 1341 cgtacacggggccactatataaacaagggtgtgcttactcagcagggagggaaataa 1400
Dh 4296 CGTACATGGGGCCATCATTAACAAGGGGCTTGCTTACCTCAGCAGGGGAATATA 4355
Qy 1401 gaacaaagagaatcttaagctatagaagccttaacttggcacaagaagctatct 1460
Dh 4356 GAACAAAGAGGAATTTCTAAGCCTATTAGAGCCGTACATTTACAAAAGGCTTACTAT 4415
Qy 1461 tatacactgtctggaacatagaagaacaaagatctcactatagagggaacagatgac 1520
Dh 4416 TATACACTGTCTGGACATAGAAAGCTAAGATCTCATATCCAGAGGAACCAAGATGGC 4475
Qy 1521 tgacgggtgtgcaagcagcagccagcgtgttaaacttgcctataatagaacggc 1580
Dh 4476 TGACCGGGTGGCCAGAGGACACCCAGGGGTGTTAACTTTCCTCATTAATAAATGTC 4535
Qy 1581 caaagcccccgaagcccaagacagctacacccagaagactggcagaagataaagaag 1640
Dh 4536 CAAGGCCCGAAGCCAGAGACAGATACACCTTAGAAGATGGCAAGATATAAAGAT 4595
Qy 1641 agaccagttcttgagactccgagggagactgctaatacctataggaagaagaaact 1700
Dh 4596 AGACA- TTCTCTGAGACTCCGGAAGAGACTCTATACCTCAGATGGGAAGAAATGCT 4654
Qy 1701 gcccccaagaaggttgaatatgtccaac- agatatacgtgttaacccactggaa 1759
Dh 4655 GCCCCAAAGAGGGTTAAATATGTCCAAAGATACATGCTTAAACCCACTGAGAA 4714
Qy 1760 ctaaacacccgcaagcttgctagaacatcccttatacgtcttgcagcctaagag 1819
Dh 4715 CTAAACACCTGACAGCATTTGGTCAGAAACATCCCTTATCATGTTCTGAGGCTACAGGAG 4774
Qy 1820 tggctgactcgtgtgtcaaacattgtgtccctgcagctggttaatgttaactccca 1879
Dh 4775 TGGGTGACTGGTGGTCAAAATGTTGTCCTGCTCCAGCTGGTAAATGCTTAATCTTCCA 4834
Qy 1880 gaataactccaggaagaagactaagggaagcaccagcgctactggaagtgact 1939
Dh 4835 GAATGCTTCCAGGAGAGACTAAGGGAGGCCACCGAGGCTTCACTGGGAATGGACT 4894
Qy 1940 tcaactgagtaaaagcggcctaaatacggnaaacaatactatgtgttltgtlgaacact 1999
Dh 4895 TCACCTGAGTAAAGCCGGCTAAATATGGAACAATACTATGTTGTTTGTGACACCT 4954
Qy 2000 ttccaaggaatggtgtagacttactactaagaagaagacttcaaccgtggtgtcaaga 2059
Dh 4955 TTTTCAGATGGGTAGAGGCTTATCTACTAAGAAAGAGATTTCAACCCGTGTGCTAATA 5014
Qy 2060 aaatacaggggaatcttccaagattggaatacctaagtaagtaaggtcagacaatg 2119
Dh 5015 AAATACCTGGAAGAATTTTCCAAAGATTGGAATACCTAAGTAAGGTTCAGACAATG 5074
Qy 2120 gtccagcttctgtcccaagtaagtaaggaactggccaagatattgggaattgtatga 2179
Dh 5075 GTCCAGCTTTGTTGCCAGGTAAGTCAAGGACTGGCCAAAGATATTGGGATTGATTGA 5134

Qy 2180 aactgactgtgcatcacagagcccaagctcaggaacaggtagagagatgataagaaca 2239
Dh 5135 AACTGACTGTGCTATACAGAGCCCAAGCTTACGACAGGTAGAGAGATGATGAACCA 5194
Qy 2240 ttaagaagacccttaaccaaatltgaacacagagactgycatataatgattgagctctcc 2299
Dh 5195 TTAAGAGACCCCTTACATAATTGACCCGGAGACTGGCGTTAAATGATGTGATACCTGCC 5254
Qy 2300 tgccttctgtcttlttgggtgaggaacaccccttgacagattgggttgacccctatg 2359
Dh 5255 TGCCCTTGTGCTTTTGAAGGTTAGGAACACCCCTGGAATTTGGCTGACCCCTATG 5314
Qy 2360 aattctacagggagaccccccgggttggcagaatctgcttgcacatagtgtcatg 2419
Dh 5315 AATTACTCTACGGGGAGACCCCCCATGTGTAGAAATGCTTCTGTACATGTGCTGATG 5374
Qy 2420 tgcgtcttccagccttgtctctcagaagcgctcagtgagtggttgagcagcag 2479
Dh 5375 TGCTCTTCCAGCCTTGTGTTCTCTAGCTCAAGGCACTTGAAGTGGGTGAGACACGAG 5434
Qy 2480 cgtggaagcagctccggaggccctactc- - - agagagagacttgcagttccacatgct 2536
Dh 5435 CGTGAAGCAACTCCGGAGGCTACTCAGAGAGGAGACTTGCAGATCCCAATGCTT 5494
Qy 2537 tccaagcttgagatcagtcatagttgaacgcacagctgacagaacactcgagactcgt 2596
Dh 5495 TCCAGTGGGAGATTCAGTCTACGTTAGACCCACCGCTGACGAACCTCAGACGCTGGT 5554
Qy 2597 ggaaggaactatctcgtacttcttgaccacaacaaagcgctggaagtggaagaaatcc 2656
Dh 5555 GGAAAGGCCCTTATCTCTACTTTTGAACCAACACCGCTGTAATAATCCAAAGAACT 5614
Qy 2657 ccacttgatcatcatccacagctlaagccggcgccacccctccgattcgggtggaag 2716
Dh 5615 CCACCTGATCATCATCATCCACGTTAGCCGGGCGCACCTCCGATTGGGGGGAAG 5674
Qy 2717 ccgaagaagctgaaatccccccttaagcttgcctccatcgcgtgttccctactctga 2776
Dh 5675 CCGAAAGACCTGAATCCCTTAACTTTCGCTCCATCCAGCTGGTTCCTTAACTGTGCA 5734
Qy 2777 ataactcccaagtaagtaagtaagccttggacagccgaactcccaataaccctta 2836
Dh 5735 ATTAACCTCTGACACTAATAGTATGCGCATAGAGACACCTCGAATCCCATTAACCTTA 5794
Qy 2837 tctctacactggttacttacttactcgtgtaacaggtatlaataltaacagcactcaagg 2896
Dh 5795 TCTCTACCTGGTAAATTAATTAAGTCCGCGACAGGTATTATTAACAAACACTCAAGG 5854
Qy 2897 gaggctcccttggggaactggtggtcgtgaatataatgtctgcttgatgaatccct 2956
Dh 5855 GAGGCTCTTTAAGAACCTGGGCTGATATATCGTTTCCCTCAAGATCAAGTATTCCT 5914
Qy 2957 ggtctcaatgacagccacaccccccctatgactcgtgttcaaggttcttactgtgc 3016
Dh 5915 AGTCT- - - - - GACTCAACCCCAATATTCCTCATGCTCAAGGATTTTATGTTTGC 5965
Qy 3017 ccagaagccccaataatgaagaataatgtggaatccctcaagatcttcccttgcagcaa 3076
Dh 5966 CAGAGACCAACCAATTAATGAAGAAATTCGCGAAATCCCAAGATTTCTTTTGAACCA 6025
Qy 3077 tggagctgataactctcaatgaatggaattggaatgagcagactcctcagcaagaaga 3136
Dh 6026 TGGAACTGTGTAACCTCTTAATGATGATTTGGAATGGCAACCTCTCACAGGATRGG 6085
Qy 3137 gtaagtactccttltgtlaaacatctcaacaggtlaataatcaatlaataatgagcaatg 3196
Dh 6086 GTAAGTTTGTATGTCAACACTATACCGAGCTGTGACATTTAATTAACCTGACCTGG 6145
Qy 3197 agatggaagaatgtgcaacagcgggttacaagaatgtgtaagaatgaataagaatgct 3254
Dh 6146 ATTAGAACCTGAAGCCCAAGTGTCTCTTACAGCTTATGATTAATAATGAATG 6203

QY 499 ctgtgatcccggtgttggaactttagcaagcccaactctaccggttaaccagaagaa 558
|||||
Db 4016 CTGTGATCCCGGGTTTGGCACTTAGACACCCACCTCAACCCGCTAACCAAGAA 4075
QY 559 ggggattctccttggtcctcgaagacagagcatttgaatgctatcaaaagccctg 618
|||||
Db 4076 GGGGAATCTCTCGGGCTCCCTGACACACAGAGGATTTGATGCTATCAAAAAGGCCCTG 4135
QY 619 ctgagcgaactgtctctgcccctccctgaagcgttaactaaaccttacccttattgtgat 678
|||||
Db 4136 CTGAGCCGACCTGCTCTGGCCCTCCCTGACCTAATCAACCTTTACCTTTATGTGAT 4195
QY 679 gaggctaaaggagtagcccgagaggttttaaccacaccccttaagacatgtagagacct 738
|||||
Db 4196 GAGCTTAAGGAGTAGCCCGAGGAGTTTAAACCAAAACCTTAGAGCATTGAGAGAGACCT 4255
QY 739 gttgctaactctgaagaagaactgtatccctgaagcagtggttgcccctatgtctgaag 798
|||||
Db 4256 GTTGCTACTCCTCAAGAAAGCTTGATCTCTACCCAGTGGTGGCCCTATGCTTGAAG 4315
QY 799 gctatcgagctgttgccatacttggcaagagcgttgacaattgacttttggagcagaat 858
|||||
Db 4316 GCTATCGAGCTGTGGCCATCTGCTCAGAGACCTGACAATTTGACTTTGGAGCAGAAAT 4375
QY 859 ataacttaatagcccccatactgcatcttgagagaactctgtcgagcccccagacagatg 918
|||||
Db 4376 ATAACTGTAATAGCCCCCAGCAATGAGAAACATCTTGAGGACCCCAAGACCGATG 4435
QY 919 atgacaaacgcccgcatacgaactatcaaaagcctgtcttctcaagagaggtcaacttc 978
|||||
Db 4436 ATGACCAAGCCCGCATAGACCCACTATCAAAAGCTGCTTCTCAGAGAGAGGTCCTTTTC 4495
QY 979 gctccacacgctctcaaacctctgcaactcttctgcttgaagagactga tgaacagtg 1038
|||||
Db 4496 GCTCCACAGCGCTCTCAACCCCTGCCACTTCTGCTGAAGACTGATGAACCAAGT 4555
QY 1039 actatgatggcatcaactatgatattgagagagacttgggtccgcaagaccccttaagac 1098
|||||
Db 4556 ACTATGATTTGCCATCAACTATGATGAGAGAGCTGGGTCCGCAAGGACCTTTAACAGAC 4615
QY 1099 ataccgtgaactggaagaagtgtcaacttgg tcaactgaagagacagctatgtgtgtgaa 1158
|||||
Db 4616 ATACCGCTGACTGAGAAAGTGTAACTGCTTCACTGACCGAAGCACTATGTCGTGGA 4675
QY 1159 ggttaagagagatgtgtggtgcgagtggttgagcgggaacccgcagatcttgggccaagc 1218
|||||
Db 4676 GGTAAAGAGATGCTGGGGGGGGGTGTGTGAGCGGAGCCACAGATCTTGCGCACAGC 4735
QY 1219 ctgcccgaagaacttaagcgcaaaaggtgactatgagccctcaacgcaagcttggcg 1278
|||||
Db 4736 CTGCGCGAAGGAATCTTCAAGGCAAAAGGCTGAGCTCATGGCCCTCAAGCAAGCTTTGGCG 4795
QY 1279 ctgcccgaagaagaatccataaactttaacgaagacagagatgacttcttgactga 1338
|||||
Db 4796 CTGGCGAAGGAATCCATAAACATTTTACGGAACAGAGATGATCCCTTTGCGACTGCA 4855
QY 1339 caagctacagggccatctataaacaaggggtgtgcttaactcaagcagagagagaaata 1398
|||||
Db 4856 CACGTACACGGGGCCCATCTTAAACCAAGGGGCTTCTTACCTCAGCAGAGGAGGAATA 4915
QY 1399 aagaacaagaagaatcttaagccctattagaagcctcatcttgcacaaaggtagct 1458
|||||
Db 4916 AAGAAACAAAGGAATTTCTAAAGCTTATTAAGAGCCCTTACATTTGCAAAAAGGCTAGCT 4975
QY 1459 attatacaactgtctcgtgacatacagaagccaagatctatactttagagagagacaagatg 1518
|||||
Db 4976 ATTATATACACTGCTCTGGAATCAGAAAGCAAGATCTATATCTAGAGAGAACAGAG 5035
QY 1519 gctgacgggttgcagacagcagccagcgtgttaacctctgcttaataagaaag 1578
|||||
Db 5036 GCTGACCGGGTTGCCAAGCAGGAGGCCGTTAACTTCTGCTTATATAGAAACG 5095

QY 1579 cccaaagccccaagaccagaagcagtagcacaccttagaagactgycagaaga taaanaag 1638
|||||
Db 5096 CCCAAAGCCCCAGAACCCAGACAGTACACCTTAGAAGACTGGCAAGATTAATAAAG 5155
QY 1639 atagcaagttctctgaagctccggagggagacttgctataactcaatgagagaaac 1698
|||||
Db 5156 ATGACCAAGTCTCTAGACTCCGAGGGGACTGCTATATCTCATATGGAAGAAATC 5215
QY 1699 ctgcccacaagaagaaggtlaagatactgccaacagatcacgtcttaaccacctaaga 1758
|||||
Db 5216 CTGCCCCAACAAGAGGTTAGATATGTCCAAACATATCATCTGTCACCCACTAGGA 5275
QY 1759 actaaacacctgcagcagatgtgtcagaacatcccttalcattgtctcgaagctacaga 1818
|||||
Db 5276 ACTAAACACCTGACGAGCTGTGTCAGAACATCCCTTATCATGTTCTGAGCTACAGGA 5335
QY 1819 gtgagctgagtggtgtaacaacttgbtggccctgcagcgttggttaagtactccctcc 1878
|||||
Db 5336 GTGGCTGACTCGGTGCTAAACATTTGTGTGCTGACCTGCTTAATGCTAATCTTTC 5395
QY 1879 agaataccctcaagaagaagactaaagggagagccacccagcgtcacttggagtgag 1938
|||||
Db 5396 AGAATACCTCCAGGAAGAGACTTAAGGGAAGCCACCCAGGCGCTCAGTGGAGATGAGC 5455
QY 1939 ttaactgaagtaaaagccggctaaatacgaagaaacaaataltatgtgttttgaacacc 1998
|||||
Db 5456 TTCACTGAGGTAAAGCCGGCTTAATACGGAACAATATCTATGTTGTTTGTACACACC 5515
QY 1999 tttaagagtggttgaagagctatctactaaagaagagacttaacgtgtgtgctaaag 2058
|||||
Db 5516 TTTTCAGATGGGTAGAGGCTTATCTTCAATGAAGAAAGACTTCACACCGTGGTCTAAA 5575
QY 2059 aaaatctcaggaagaatlttccaagatttgaatacctaagtgtaaggttcaagacat 2118
|||||
Db 5576 AAAATCTGAGGAATTTTTCGAGATTGTGAATTAAGTATGAGGTGACACAT 5635
QY 2119 ggtcagacttctgttgcacaggttaagtcagagagcttgcacaaatattggagattgtg 2178
|||||
Db 5636 GGTCAAGCTTTTGTGCCCCAGGTAGTCAAGGAGCTGCGCAAGATTTTGGGATGATTGG 5695
QY 2179 aaactgcatctgtcatalcagaagcccaagagctcaagagagatgaatgaataaac 2238
|||||
Db 5696 AAACGTGATTTGTCTACAGACCCCAAGCTTCAAGACAGTGAAGAGATGAATGAACCC 5755
QY 2239 attaaagagaccccttaccacaaatttgacacagaagacttgcattgaattgtgactc 2298
|||||
Db 5756 ATTAAAGAGACCTTACCAAAATTGACCAAGAGATGTCATTAAATGATTGATGACTGTC 5815
QY 2299 ctgccccttgtgtctttagaggtgaagaaaccccttgacagcttgggttgaacccctat 2358
|||||
Db 5816 CTGCCCTTGTGCTTTTAAAGGTTAGGAACACCCCTGACAGTGTGGCTGACCCCTAT 5875
QY 2359 gaattgctctagggagagcccccctgttgacagaatltgcttgcacataagtgat 2418
|||||
Db 5876 GAATTGCTCTAGGGGGACCCCCCGTTGTAGAAATTTGCTTGTATCATATGTCGTAT 5935
QY 2419 gttcgtcttcccaagccttctctcctaggtcgaagggcgtgagtgagagcaga 2478
|||||
Db 5936 GTGCGCTTTCCACACCTGTTCTCTAGGCTCAAGAGCGCTCGAGTGGGTAGGACAGA 5995
QY 2479 gcttgaagcagctcccgagagcactactcaag ---aggagacttgcagttccacatgc 2535
|||||
Db 5996 CGGTGAGAGAGCTCCGGAGGAGCTCAAGAGAGAGAGACTTGCAAGTTCCACATCGC 6055
QY 2536 ttccaagtltgaagatcaagctatggttagagcagcgttgcaggaacacctcgagactcg 2595
|||||
Db 6056 TTCCAAGTGGAGATTCATGTTATGTTAGAGCCACCGCTGAGGAAACCTGAGACTCGG 6115
QY 2596 ttgaagagaccttctcgttacttgaacacacacacacacacacacacacacacacacac 2655
|||||
Db 6116 TGGAAAGGCGCTTCTCTGATCTTTTGAACACACCAAGGCGTGTGAAGTGAAGGATC 6175
QY 2656 cccacctggaatcgaatcccaagctaaagccggcagcactcccgatctcgagg tggaaa 2715

```

|||||
Db 6176 TCACCTGATCATACATCCACGTTAAGCTGGGCCACCTCCGAGCTGGGGTGGAGA 6235
Qy 2716 gccgaagaactgaaatcccttaagctgcctccatccgctggttccctactctgtc 2775
Db 6236 GCCGAAAGAACTGAGAAATCCCTTAAAGCTTCCCTCCATCGCCTGGTCTTACTTAC 6295
Qy 2776 aataactcctcaagttaagccttctgtgacagccgaatcccaataaaccctt 2835
Db 6296 AATAACTCCCGACGCGCAGTAGTAACGCTTATAGACAGCTGAAACCCCATAGACCTTT 6355
Qy 2836 atctcaccctgttactactactcgcgtacaggtatataatatacaagcactcaag 2895
Db 6356 ATCCCTTACTGCTGCTATTATTACCCCTGATACGGGTGTCACGTAAATAGACCTGAG 6415
Qy 2896 ggaagctcccttgggagcctgtgtgcctgaatataatctctccttgcctgaatccc 2955
Db 6416 TGTTCCTCTAGAGGCACTGGGCGCTGACACTGCTTCTGCTCCGATTGATTAAACC 6475
Qy 2956 tggctcaatgaccagcagcaccaccccgatgactcgttctgaaggttactgctt 3015
Db 6476 CGCTGTTA-----AAAGCACACCTCCACACTAGTCCGATTATGAGGTTCTATTGCTG 6529
Qy 3016 cccaggaaccacaaatagaatataatgtggaatccctcaggaattccttgcagaag 3075
Db 6530 CCCAGG---CACAGAGAAAGAGAAATACGTGGGGAATCTGGGGGAATCTCTGTAGAG 6586
Qy 3076 atgagactgcataacttctaataatggaatgtggaatgcagctctcagcaagacag 3135
Db 6587 ATGAGAGCTGCTGCTACCTCCACAGCTGAGAGTGGAAATGGCGGATCTCTCCAGAGACG 6646
Qy 3136 agtaagtaactcttctgttaa 3156
Db 6647 GGTAAATTTCTCTTTGTCAA 6667

RESULT 13
PEN133818
LOCUS 8750 bp DNA VRL 14-APR-2000
DEFINITION Porcine endogenous retrovirus type C proviral gag, pol and env
genes and LTR (class B, clone 43).
ACCESSION AJ133818
VERSION AJ133818.1 GI:6688949
KEYWORDS env gene; env protein; gag gene; gag protein; pol gene; pol
protein.
SOURCE
ORGANISM
porcine endogenous retrovirus.
porcine endogenous retrovirus
viruses; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
1 (bases 1 to 8750)
REFERENCE
AUTHORS Czaderna,F., Fischer,N., Bolter,K., Kurth,R. and Tonjes,R.R.
TITLE Establishment and characterization of molecular clones of porcine
endogenous retroviruses replicating on human cells
JOURNAL J. Virol. 74 (9), 4028-4038 (2000)
2 (bases 1 to 8750)
MEDLINE 20219394
REFERENCE Toenjes,R.R.
AUTHORS Direct Submission
TITLE Submitted (04-MAY-1999) Toenjes R.R., Medical Biotechnology,
JOURNAL Paul-Ehrlich-Institut, Paul-Ehrlich-Straasse 51-59, Hessen, D-63225
Langen, GERMANY
FEATURES
location/Qualifiers
source 1..8750
/organism="porcine endogenous retrovirus"
/proviral
/strain="type C"
/db_xref="taxon:61673"
/clone="43"
1..629
LTR 467
misc_feature /note="cap site"
primer_bind 632..649
/note="gly4"

```

gene
mat_peptide

CDS

```

1064..2635
/gene="gag"
/gene="gag"
1064..2635
mat_peptide
/function="retroviral capsid"
/product="Gag protein"
join(1064..2635,2639..6223,6099..8072)
/note="Viral polypeptide"
/codon_start=1
/protein_id="CAB65341.1"
/db_xref="GI:6688950"
/translation="MGQVTTPSLIDHTVRSRAHNLISVYKGGPNOTPCASEMP
TFDWGSESTFENSEILLAVKAILPQPSHSDQERYLLTWDLDEDPWPKWMLN
KPRKPGRIIALGKMKNSAKKVEPSRIPELIEPTPEPPEPPVPAQAVG
SPAPGAPVEGPACTRRSRRAATPEETDEIALLPTTPYPPMGGLOLQYAPSS
ADLYNTNHPSEDEPRLGLVESIMESHOPMDCCOOLQTLFTTERERLILEA
RKNVPGADKPTQLOQLDIMGFLTRGCMGYNTAEGESIKITROLAVAGLRASRP
TNLAKVEVMQGPNEPVSFLERLMEAFRRFTTPDPISSEAKSVALATIGOSALDR
KILRLGLQEAELDLVRAEKVYIRRETEEEKERERERERERERERERERER
KILAAVEGKSSREREDFRKIRSGPSONLGNRPDLDDQACYCKGMAKNCCK
KNGKPKVLALIEDKDRGSDPLPEBRVALKEGQVEFLVDGAEHSLQPLGLK
KEKSMVMGATGOROYPTTTRRYDLAVGHVTSFLYIPCPVPLGLRDLITKMAOI
SEBGRREVSNKPIITVLLTDLDEFLYSPQVKPDODIOSMLDOPRMAEAGAG
LAKVPPQVILKASATPVSYROTPLSRKARBGIMHYOGLIIVPOSPMPTL
LVRKPGTNDYRPVODLREVKRRODILHPVNPYNLSLPEBNMTYVLDKDAFT
CLRHPPTSOPLEAFEMRDPOTGRGQTLWRLPDGFKNSTPIFDEALHRLANFRIOH
POVTLQYVDILLAGATKODCEGTAKALLLESDLGYSARAKAQICREVVLYGYS
LGGORMLTEARKKTVVOIPAPPTAKQVREELGAGFCRLMIGPATLADPLYTE
KGEFSNAPBHOCTPDATIKKALISAPALADVYKRPFLVYDERGVARGVLTQIGPW
RRPAYSLSKLDIVASGMPVCLKIAVAAILVDADKLTIGONTIYAIPALBNRYVO
PDDRMNTNARNTHYOSLLLEVTFAFPALINPTLLPEETDEVTIDHQLILEBTG
VRKDLDTPLTEGEVLWTFEYSSVVGKRMAGAAYVDGHTTMASSLPGTSAOKAE
LMAITQALRLAEKSIINITYDSRYAFTAVHGAIRKQGLITSAGREINKEEISTL
LEALHLPRLAIIHCPGHOKADLISGNOMADVAKOAAVALLPIETPKAPEPR
ROYTLEDMOEIKITDQSETEPEGCTSYSGKEILPHKEGLIYVOQLIRHFLGKHLQ
QVARTSPHVLRLPGVADSVYKCPQOLVNNANSRYPCKRRLGSRGSHAMVDFTE
VPRKAYGNKYLLVFDVTFSGWEAVPYKKESTVAVAKLIEETPRGILKIVGSDNG
PAPVAYROSGIYLLIDMKLHCAYRROSSQOYERMMKTIKETLTKLTTEGINDVLA
LLPEVLFVRNTPGFGLTEYELLYGDPPLVEIASVSHADVLISQLEFRLALAEV
RORAMKOLREAYSGEGLQVDFHREOVDSYVRRHRAAGNLETWKGPYLLVLTPTAV
KVEGISTMIAHSYKLAIPPDGSGRAKTEPKLRHLRLHLYVSNNSPOMPTLSM
RHLPTRGGERKRLKIPISFASIANFLTLITROASRKLIDSSMHPRLITLILDP
DTGVTNSTRGVAPRGTWPELHPCLLINPAVSTPNIRSYGTCCTGTEREKYC
GSESEPCRRMSCVTSNDGDKWPISLQDRKFSVYSGPKYVMKLVYDKSCSPSD
LDYLKISFTEKQENIQKWMGSMGIVFYKGGAGSTGLITLRLETSDEPVAAG
GKVLAEQGPALPEPNNLPEVPLTSRPDITOPPSGTGLITLRNPRSPGVAVT
GORLPSLIGAFQAINSTDPATSCMLCSSGPYEGMAKEGFWTKRNMOCNTL
GSRNKLTITEVSGGTGCGKAPRSHOULCYSTVYEDASENOXYLPGVNNMNCNTL
TPCVSSVFNSKDPVWQVIVPVYTHPREPVYLDVDEYDIRNRKRPVSLTAVMIG
LGTAVGVTGTAALITGPQOLEKGLGELHAMTEDLALAEVSANLESLSLSEVVL
ONRGLDLIFLREGGLCAALKKECCFYVDHSGALRDSMSKLRELERRREREDQGM
FKGMFNRSPMTTLLSALTGLVLLVLLLVGCLINRFVAFVERVSAVOIMVLRQO
YGLISGSETDL"
2639..6223
/gene="pol"
/function="retroviral protease and polymerase"
/product="Pol protein"
2639..6223
/gene="pol"
/gene="pol"
6099..8069
/gene="env"
/function="retroviral envelope"
/product="Env protein"
6099..8069
/gene="env"
/gene="env"
1..8750
LTR POLYA-signal
BASE COUNT 2288 a 2286 g 1962 t
ORIGIN

```

Query Match 86.3%; Score 2864.6; DB 14; Length 8750;
Best Local Similarity 95.2%; Pred. No. 0;

QY	19	acaagtgctcttcttgctgcctgaagtacacccaactagccaacaacttttgcttcga	78
Db	3446	AAAGATGGCTTCTTCTTGCTTAGATTAACACCCACATAGCCACCCCTTTTGCTTCGA	3505
QY	79	tgaagaataccaagtaacggaagaaccggacatacaactcggacccgactgcaccaagg	138
Db	3506	TGGAGAGATCCAGGTACGGGAGAACCGGGCAGCTACCTGGACCCGACTGGCCCAAGG	3565
QY	139	ltcaagaactcccgacacatctttgaagaaagccctacacagaagacctgtgcaacttcag	198
Db	3566	TTCAAGAACTCCCGACACATCTTTGAGAGAGCCCTACACAGGAACTGGCCAACTTCAG	3622
QY	199	atccaacaccttaagtgtagccctccctccagtaagtgatgaactgtcttgccygaagcc	258
Db	3626	ATCCAAACCTCTCAGGTAGACCTCTCTCCATAGCTGATGATGACTGCTTGTGGCGGAGCC	3685
QY	259	accaaacagagactcttagaaggtacagaagacactctgtcgaattgtctgaactagc	318
Db	3686	ACCAACACAGACTCTTTAGAAAGTAGAAGGACACTACTCTGTGAATTGTCTGACTTAGCC	3745
QY	319	tacagagcctctgtctaaagaagcccaagatttgcaagaagagtaataactcttggtgac	378
Db	3746	TACAGAGCCTCTGCTTAGAAGGCGCCAGATTTCGAGAGAGAGCTAACACTTGGGGTNC	3805
QY	379	agtttgcygggcygcygaagcgaatgctgtacgcygaacggaagaaactgtactccaata	438
Db	3806	AGTTTGGGGGGCGGGCGAGCATGGCTGACGGAGGACAGGAGAAACTGTAGTCAGATA	3865
QY	439	ccggcccccaacacagccaaacagatgagaagatttttgggaacagctgtgaatttgcga	498
Db	3866	CCGGCCCCAACACAGCCAAACAGTAGAGATTTTGGGAGACGCTGGATTTTGGAGA	3925
QY	499	ctgtgtgaatcccgaggttttgacacttagacagcccaactaacccgtacaagaagaaaa	556
Db	3926	CTGTGTGATCCCGGGGTTTGGACCTTAGAGAGCCCACTTACCCCTTACCAAGAAATA	3985
QY	559	gggggattctcctgtggtcctctgaagaccagaagcatttgatgtatlaaanaagccctg	618
Db	3986	GGGGAAATTCCTCCGGGCTCTTAGACACAGAAAGACATTTGATGCTATCAAAAAGGCCCTG	4045
QY	619	cttagcgacactgtctctgtgacctccctgaacgtaactaaaccttacccttatgtgat	678
Db	4046	CTTAGCGCACCTgtctctgtgacctccctgaacgtaactaaaccttacccttatgtgat	4105
QY	679	gaagcgtaaaggagtagcccgagagattttaaaccsaacccttagacatatgagagaaact	738
Db	4106	GAGCGTAAGGAGTAGCCCGAGAGATTTTAAACCAAACTTAGACATAGGAGAGACCT	4165
QY	739	gttgctactctgtcaagaagctgtatcctgttagccagtgtgttgcccgatgtctgaag	798
Db	4166	GTTGGCTTACTGTCAAAAGAGCTTGTATCTGTAGCGAGTGTTGGCCGTATGCTTGAG	4225
QY	799	gctatcgagagctgtggcatatctgttcaagaagcgtcaaaattgactttggagagaat	858
Db	4226	GCTATCGACGCTgtgccaatCTGCTCAAGAGACCTTACAATTTGACTTTTGGAGAGAAT	4285
QY	859	ataactgtlaatgcccccaatgtcatgttgaagacatgttggagcccccaagcgaatg	918
Db	4286	ATAACTGTATATGCCCCCATGTGATTTGGAGAACATCTGTTGGCAAGCCCCCAAGCCGATG	4345
QY	919	atgaccaagcccgcatgataccaactatcaaaagcctgtcttctcaagagagtgtaacttc	978
Db	4346	ATGACCAACAGCCCGCATGATGCCCATATCAAAAGCTGCTTCTACAGAGAGGGTCACTTTC	4405
QY	979	gttcacacagccgcttcaaaccttgcaactcttctgtccttgaagagactatgaacagt	1038
Db	4406	GCTTCACAGCGGCTCTCAACCTTGCCCACTTCTGTGCTGAAGAGCTGATCAACAGTG	4465
QY	1039	actcaatgtgcatacaactcttattatgagagagactgggtccgaaagacttataagac	1098
Db	4466	ACTCATGTATGCTATCACTTTTATTTAGAGAGACTGTGGGTCCGCAAGAGACTTTACAGAC	4525

QY	1099	ataccgcgtacgcgtggaagctgtcaaccctggtttcacttaacggaacgactatgtytgtaa	1158
Db	4526	ATACGCGTGACTCGGAGAACTGTAACCTGGTTCACTGACGGAAGCAAGCTATGCTGTGAA	4585
QY	1159	ggtlaagagatggtctggygcgcaatggtgtagcagggaccgcacagatctggtccagcgc	1218
Db	4586	GGTAAAGAGATGCTGTGGCGGGGTGTGTGACGGGACCACACAGATCTGGGCCACGACG	4645
QY	1219	ctgcccgaaggaacttcaagccgaaggctgagctcatgacctcaogcaagcttgcg	1278
Db	4646	CTGCCGGAGGAACCTTCAGCCCAAAAGGCTTAGCTCATGCGCTTCACGCAAGCTTTGCGG	4705
QY	1279	ctgcccgaaggaatccataacatttatacgaacagcaggtatgtgctttgacgtgca	1338
Db	4706	CTGCCGGAAGGGAATCCATTAACAATTTATAGGACACGAGATGATGGCTTTGGCCACGTCA	4765
QY	1339	caagtlacacggggccattataaaacaaagggtgtgtcttaccctcagcagggagagaaata	1398
Db	4766	CACGTACACGGGGCCATCTATTAACAAGGGGTTGCTTACCTACGACGAGGAGGAATA	4825
QY	1399	aagaaacaaagggaattctaaagcctattagaagccttaccatltgcccanaaagctagct	1458
Db	4826	AAGAACAAGGGAATTTCAAGCCTATTGAAACCTTCATATTGGCCAAAAGGCTAGCT	4885
QY	1459	attataacgtcttcctgagactccaagaagccaagaatctctatctagaaggaccagatg	1518
Db	4886	ATTATACACTGTCTCGACATCTGAAAAGCCAAAGATCTCATATCTAAGAGGAACCGATG	4945
QY	1519	gctgaccgggtltgcgaagcagcccaagcgtgttaacctctgctcttaatatagaagc	1578
Db	4946	GCTGACCGGGTGGCCAGCAGGACGCCAGGCTTAACTTCTGCTTAAATAGAAAG	5005
QY	1579	cccaaaagcccccagaaccagaagacagtlacacctagaagactgtgcaagagataaaag	1638
Db	5006	CCCAAAAGCCCCAAGACCCACACACAGCATGACCCCTGAAGACATGCGCAAGATMAAAAG	5065
QY	1639	atagaccagctctctgaagctccggaggagacctgtctataccctcaatatygaagaatc	1698
Db	5066	ATGAGCCAGTCTCTGTGAAGTCCGGAGGGACCTCGTTAAACCTCATATGTGGAGGAAATC	5125
QY	1699	ctgcccacaaagaagggttagaatatgtccaacagatacatcgcttaaccacactagga	1758
Db	5126	CTGCCCCACAAAGAAGGTTAGATATGTGTCCAAAGATACATCTGTTCAACCACTTAGGA	5185
QY	1759	actaaacacctgcagcagttgtgcagaacatcccttalcattgtcttgagcttacagga	1818
Db	5186	ACTAAACACCTGCAGCAGTGTGGTCACAAACATCCCTTATCATGTGTGAGCTTACAGGA	5245
QY	1819	gttgccagactggtgtgtcaaatatgtgtgtccctgcacagctgtgttaatgtcaatccctcc	1878
Db	5246	GTGGCTGACTCGGTGTCTAAACATTTGTGTGGCTCTGCACCTGTTAAATGCTAATCTTCC	5305
QY	1879	agaataccctccagaaagagactaaagggaagccaaccacagcgcacatctgggaagtgaac	1938
Db	5306	AGAGTACCTCCAGGAAGAGACTTAAGGGGAACCCACCAAGGCGTCACTGGGAATGTGAC	5365
QY	1939	ttaactgaggtaaagcccggttaataacagaaacaaataltatltgttlttltgtagaac	1998
Db	5366	TTTCACTGAGGTAAACCGCGCTAAATACGAAACAAATATATTTGTTTGTAGACACC	5425
QY	1999	ttttaagatgtgtgtagagcttatctctactaagaagaagactcaaccgtgtgtgctaa	2058
Db	5426	TTTTTAAGATGGGTGAGAGCTTATCTCTAAGAAAGAGACTTCAACCGTGTGTGCTTAA	5485
QY	2059	aaaataactggaagaaattttccaagacttltgaatatacctaaggttaataggctagaacat	2118
Db	5486	AAAATACGTGGAGGAATTTTCCGAGATTTGGATTAAGGTAAATGCGGTGTGACACAT	5545
QY	2119	ggtccagcttcgttgcaccagtaagtcagagactgtgcagaatatatvggagatgtatg	2178
Db	5546	GGTCCAGCTTTTGTGCCACAGTAAATGACGAGACTGGCCAAAGATATTTGGGATTTGATGG	5605

QY 2179 aaactgattgtgatacagacaccccaagctcagagcaggtagagagatgaaatagaacc 2238
 Db 5606 AAATGATTTGTGATACAGACCCCAAGCTCAGAGACAGTACAGAGATGAAATAAAC 5665
 QY 2239 attaaagagaccccttaccatattgacacagagacgtgacattatgattgagtcctc 2298
 Db 5666 ATTAAAGAGACCCCTTACCAATATGACACAGAGACGTGGCATTAATGATGATGCTCTC 5725
 QY 2299 ctgaccccttgctctttagaggtgagaaacaccccttgagacacattgagccctat 2358
 Db 5726 CTGCGCTTTGCTTTTATGAGGTTAGAAACACCCCTGAGAGTTGGGCTGACCCCTAT 5785
 QY 2359 gaattgctcactagcgggagaccccccgttgcagaaattgctcttcacatagtctgcat 2418
 Db 5786 GAATTGCTCTAGCGGGAGACCCCGCTGTGTAATTCCTCTGTACTGATGCTGAT 5845
 QY 2419 gtgtgttttccagccttgtctctcctcctcagagcgtcgtgagtggtgagcagcga 2478
 Db 5846 GTGCTGCTTCCAGCGCTGTTCTCTAGGCTCAGAGCGCTCGAGTGGGAGCAACGA 5905
 QY 2479 gctgtgagacagctcgcggagcctactcag--aggaaacttgaaattccacatgc 2535
 Db 5906 GCGTGAAGACAGCTCCGGGAGGCTTACTCAGGAGAGAGACTTGCAAGTTCCACATCG 5965
 QY 2536 ttccaagltggagattcagctcactatgttagacgcacacgtgcaggaacctcagactcgg 2595
 Db 5966 TTCCAAGTGGGAGATTGAGTCTATGTTAGACGCCACCGTCCAGAGAAACCTCGAGACTCG 6025
 QY 2596 tggagagagactatctcgtacttttgacacacacacagcgtgtgaaagtcgaagaaac 2655
 Db 6026 TGGAGAGGCGCTTATCTGACTTTTGACACACCAAGCGCTGTGAAGTCGAAGCAATC 6085
 QY 2656 cccacccgtgatacagatcccaagcttaagcgcgcacacccctccagctcgggtgagaa 2715
 Db 6086 TCCACCTGGATTCATGATCCACGTTAACTGGCCGACCTCCACATCGGGTGAAGA 6145
 QY 2716 gccgaaaaagctgaaaaatcccttaagctcgcctccatgcgtgttcttactctgctc 2775
 Db 6146 GCGGAAAAAGCTGAGATCCCTTAAGCTCGCTCCATCGCGCTGCTTACTCTAAC 6205
 QY 2776 aataacccctcaagttatgttaagcgtctgtgagacgcgcgaacctcccataaacct 2835
 Db 6206 AATAACTCCCAAGGCGCAGTAGTAACGCTTATAGACAGCTCGAACCCCATAGACCTTT 6265
 QY 2836 attctcaactgttactactcctaccccgtagacaglatataatataacagacactaaag 2895
 Db 6266 ATCCCTTACTGGCTGATTTATGACCTGATACGGGTGCACGTAAATAGCACTGAGAG 6325
 QY 2896 gtaggctcccttggagacgtgtgcctgaatatagttcgctcctcgaatcaatcc 2955
 Db 6326 TGTGTCCTTAGAGGACAGCTGGTGGCTGAGACTGATTTCTGCTCGATTGATTAACC 6385
 QY 2956 tggttcaatgacacagccacaccccgatgactccgtgtccttaagggttttaagtttg 3015
 Db 6386 CGCTGTTA-----AAAGCACACCTCCCAACTAGTCCGATGATGAGGTTCTATTGCTG 6439
 QY 3016 cccagagaccccaatataatgaagatattggaatccctcagagattccttgcagaaga 3075
 Db 6440 CCCAG--CACAGAGAGAGAGAAATACTGTGGGGTCTCTGAGGAATCTTCTGTAGGAG 6496
 QY 3076 atgagagctgataactcctaattgagtggaattggaaatgagcagctcctcagaagaagag 3135
 Db 6497 ATGAGAGCTGCTGCTCCTCCAAAGCTGAGAGACTGGAATGAGCGATCTCTCCAGGACCG 6556
 QY 3136 agtaagtaactcttctgttaa 3156
 Db 6557 GGTAATAATCTCCTTGTCA 6577

RESULT 14
 LOCUS PEN293657 8763 bp DNA VRL 30-MAY-2001
 DEFINITION Porcine endogenous type C retrovirus proviral gag gene, pol gene

and env gene, class B, clone 213.
 AJ293657
 VERSION AJ293657.1 GI:14275841
 KEYWORDS env gene; envelope; gag gene; group specific antigen; pol gene; polymerase; polyprotein.
 SOURCE porcine endogenous type C retrovirus.
 ORGANISM porcine endogenous type C retrovirus.
 REFERENCE 1 (bases 1 to 8763)
 AUTHORS Krach, D., Fischer, N., Czauderna, F. and Tonjes, R.R.
 TITLE Comparison of replication-competent molecular clones of porcine endogenous retrovirus class a and class b derived from pig and human cells
 JOURNAL J. Virol. 75 (12), 5465-5472 (2001)
 MEDLINE 21256017
 REFERENCE 2 (bases 1 to 8763)
 AUTHORS Tonjes, R.R.
 TITLE Direct Submision
 JOURNAL Submitted (31-JUL-2000) Tonjes R.R., Medical Biotechnology, Paul-Ehrlich-Institut, Paul-Ehrlich-Str. 51-95, Langen, 63225, GERMANY
 FEATURES
 source
 Location/Qualifiers
 1..8763
 /organism="porcine endogenous type C retrovirus"
 /proviral
 /specific_host="Sus scrofa"
 /db_xref="taxon:70540"
 /clone="213"
 /cell_line="PK15"
 /note="class B"
 1..630
 467
 /note="cap site"
 633..650
 1077..2651
 /gene="gag"
 1077..2651
 /gene="gag"
 /codon_start=1
 /product="group specific antigen"
 /protein_id="CAC39625.1"
 /db_xref="GI:14275842"
 /translation="MGOTVTPPLSLTLDHWTEVRSRAHNLISVQKKGPMOTFCASEMP
 TFDVMPSEGFENSEILAVKAILFQPGSHPDPEYITLWDLAPEPPVWPKMLN
 KPRKPSRILALGKNSHAKSPSGCIPEPEPTMPEDVPPVPAAGAVG
 PSAPKAPVYBGPAGTRSRGATPEPTDELAIIPLTTYPPMGGOLQPLQTPSS
 ADLNNKTNHPPESDEDQRLGLVESLMSFHPWDDQQLDQLFTTERERILLLEA
 RKNVGDADGRPTOLQNEIDMGFLTRGMDYNPAGERESLKIYQALVAGLRASRP
 TNLAKVEVMQGPNEPSPFLERIMEAFRRFTPDPTSEAKASVALAFIGSALDIR
 KILQRLGLOEALRDLVREAEKVIYRRETEEEKREKEREEREERERDRROENLT
 KILAAVEGSSSRERERDFKIRSGPROSONLGNRTPLDNDQACAYCKEKHMANRNCX
 KGNKGPVIALIEDKD"
 2652..6239
 /gene="pol"
 <2652..6239
 /gene="pol"
 /note="complete pol ORF, gag-pol is expressed as a polyprotein"
 /citation=[1]
 /codon_start=1
 /product="polymerase"
 /protein_id="CAC39626.1"
 /db_xref="GI:14275843"
 /translation="GRGSDPLPEPRITLAVEGQVPEFLVDTAHSHVLLQPLGKLKE
 KSNWKGATGROTPMTTRKTVDLAVGRVTHSLVPECPVLIGRLILTKMAQISF
 EQGRPEVSNKPLITVLLQIDDEYRLYSQVVPDDIOSMLQPOVAMEVGMGLA
 KOYPPQVTLKASATPVSROYPLSREARGLIMPHVQRLIOGILVAVSPPMTPLP
 VRKPGTNDYRPVODLREVNRVODIHPITVNPYLIALPPEERNWTVLDLKAFFCL
 RIHPTSQLFAFERMDPGTGTGQLTTRLRDQPKNSPFFEDALRDLANPFIQHPQ
 VTLLQYVDLLLAGATQDCECTKALLLELSLGLRASKAKKQICREPTTYIGYSLR
 GQQRWLLBARKTVVQILPPTTAKQVREFLGTAGFCFLWPGFATLAAPLYPLTKRG
 EFSWAPHQKAFDAIKKALLAPALAPVTKPFTTYLVDNRKGVAGVGLVDTGLGPMRG
 PVALTSKKLDPVASGMPVCLKALIAAVALIKVDADKLLTIGNITIVIAPIHALENIVRQPP

DRMTNARMTHTYOSLLTERTVFPAPALNPATLPEETDEVTYDCHOLIEETGVR
 KDLDIPLTEGVLTPFTDSSSYVEEGKAGAAVNDGFTIASSLPEETSQAKELM
 ALIOLARLAEKGSINITYDSRYAFATAHVAHYIYORGLITSAEIKKEILILE
 ALHLEPRLAIEHCPGHOAKADLISRONOMADRVANAOAVMLPIETPKPEERRO
 YLLEDMOEIKIIDOFSEPEEGTCLISYKEILIPHEKELEYOUIHRLHGTKHLOOL
 VRTSPYHVALRPGVADSVYKHCVPOCIANVANSRIPOGRKIEISPGANWEDPEVVK
 PAKYKNKYLIVYEDTFSWVEAYPTKKESTYVAKKILBEITPRGIRKIVISDNGPA
 EVAQVSQGLAKLIGIDMKLCAVRPOSSQVEERMNTIKETLTKLTETGINDMIALP
 PVLFRVNTPOGFLPYPELLYGGPPPLVETASVADVLISQPLFRSLKALEVRO
 RAMKOLREAYSEEGDLOVHFQVDSVYVRHHAGNLETFRKGYLVLTLTPYAVK
 EGISTWISHASHKVLAPPDSCWRKAEKTEMLRLRLHPLVPSYNNNSPQ"

gene
 6112..8085
 /gene="env"
 6112..8085
 /gene="env"
 /codon_start=1
 /product="envelope"
 /protein_id="CAC39627.1"
 /db_xref="GI:14275844"

CDS
 /translation="MHPTLSMRHLPTRGGEPRRLIPLSFASIAWFLTLTTPQASSK
 RLIDSNPHRPLSLTWLIDPDTGVNTSGVARGTMMPELHCLILINAVKSTP
 PNLVRSYGCPTGEKEKCYCGSGESFCRHSCTVSNIDGDKMPLISODRKFSEVN
 SGEGTKVMKLYKDKSCSPDLDYLKISPEKQENIOKMGMSGLIVEKYGGA
 GSTLRLRLIEGTCTPPVAPGPKYLAQOGPALREPHNLPPVPLTSTLRDTPPSX
 STTGCLPTNPNSPCVYKTOQLSLIQGALQAINSDPDATSSCHMLCUSDGPPY
 EEMAKEGKFNVTKEHNDCTWGSNNKLTLEVSQGTGIGKAPSHOHKESTVVEQ
 ASENDLVPNGVNRMAKNTGLPCVSTVFNSQKVOQVLPVRYHPEVVDVEY
 DYRNPKEPEVSLTAVMLGTAAGVGTALITGPVQOLEKELGELHSMGATDLR
 ALEESYNSLESLTSEVILONRGDLTLFREGGLAALKKECCFVVDHSAIRDS
 MSKIRLERLRERREREDOGMEGFNPSMTTLTSLTGLVYVLLLLTLVPGCLIN
 RFVAFVREVSVAQIMVLRQYGLISQETDL"

LTR
 8133..8763
 8655..8660
 POLYA_signal
 2301 a 2200 c 2287 g 1974 t 1 others
 BASE COUNT
 ORIGIN

Query Match 86.3%; Score 2864.6; DB 14; Length 8763;
 Best Local Similarity 95.2%; Pred. No. 0;
 Matches 2990; Conservative 0; Mismatches 139; Indels 12; Gaps 3;

QY 19 acagatgacctctctgacctgaattacacccactagccaacacttttgccttcga 78
 Db 3459 AAGAGTCCCTTCTTCTGCGGAGATTACACCCCACTAGCCAACTTTTGGCTTCGA 3518
 QY 79 tggagagatccaggtacggaagacccgagctacacctgagaccgagcccaag 138
 Db 3519 TGGAGAGATCCAGGTACGGGAAGAACCCGGGACACTGACCGGACCTGGCCCAAGG 3578
 QY 139 ttcaagaactccccgacatctttgacgaagccctacagaagacctgycgaactcaag 198
 Db 3579 TTCAAGAACTCCCGAATTCTTTGACGAAGCCCTACACAGGAGACTGGCCAACTTCAAG 3638
 QY 199 atccaacacctcaggtgacctccctccagtaagtgagacctgtcttgccggagacc 258
 Db 3639 ATCCAAACCCCTAGGTGACCCCTCCCACTAGCTGATGACCTGTTTGCGGGAGACC 3698
 QY 259 accaacaagagctgcttagaaggtacgaagcactactgctggaattgctcgaactaagc 318
 Db 3699 ACCAAACAGAGACTGTTAGAGGTACGAAGGCACTACTGCTGGAATTGCTGACCTAGGC 3758
 QY 319 tacaagacctctgcttagaagagccagatttgagaagagaaggttaacataacttgggtac 378
 Db 3759 TACAGAGCTCTGCTAAGAGGCCAGATTGAGAGAGAGAGGTAACTTGGGGGTAC 3818
 QY 379 agtttcgagggcgagcagatgctgacgagagcaggaagaaacttaagtcagata 438
 Db 3819 AGTTTGGGGGCGGAGGATGCTGACGAGGAGCGAGGAAGAACTTATGTCAGATA 3878
 QY 439 cgggcccccaacacacaaagtgagagagttttggggacagctggaatttgcaga 498
 Db 3879 CCGGCCCAACACACAAAGTGAGAGAGTTTGGGGAACAGCTGATTTGGACA 3938

QY 499 ctgtgattccgggggttgagaccttagcagccccacttaaccgcttaaccagaanaa 558
 Db 3939 CTGTGATCCCGGGGTTTGCGACCTTAGCAGCCCACTACCCGCTTAACCAAGAAAA 3998
 QY 559 ggggagattccttgagcccttagacacagaagagatttgatcatataaagccctg 618
 Db 3999 GGGGAATTCCTCTGGGCTCTGAGCACCAGAAAGGATTGATGCTATCAAAAGCCCTG 4058
 QY 619 ctgaacgcaaccctgaccttgacctctgaagtaacttaacccttaacccttaatgag 678
 Db 4059 CTGACGACACCGCTGCTGACCATCTGACGTAACCACTTAACCTTTATGTGAT 4118
 QY 679 gacgctaaaggagtagcccgagagtttaaccaccaacccttagaagcaatggaagacct 738
 Db 4119 GAGCCTAAGGAGAGTAGCCCGAGAGATTGTAACCCAAACCTTAGGACCATGAGAGAACT 4178
 QY 739 gttgacctctgctaaagaagcttgatctcttagcagaatggttgccggtatgctgaag 798
 Db 4179 GTTGCCTACCTGTCAAAGAACCTTGATCCTTAGCCAGTGTGGCCCGTATGCTTGAG 4238
 QY 799 gctatcgagctgtgacatactggtcaaggagcgtgacaacttgactttgagacagat 858
 Db 4239 GCTATCGGACGCTGTGACCATCTGCTCAAGAGCCTGACCAATTTGACTTTGGACAGAT 4298
 QY 859 ataactgtaataagcccccaatgcatltggaacacatcgltggcagccccagacgat 918
 Db 4299 ATAACCTGAATAGCCCCCATGATGAGGAACATCTGTCGGCAGCCCGCCAGACGATG 4358
 QY 919 atgacaaagcccgagatcacacactacaagacgtgcttccaaagagaagtgacttc 978
 Db 4359 ATGACAAAGCCCGATACCCACTATCAAGACCTGCTTCTCAAGAGAGGCTACCTTC 4418
 QY 979 gctccacagcgcgtctctaacaccttgccactctctctgacgaagagactgataccagtg 1038
 Db 4419 GCTCCACAGCGCGCTCTCAACCCCTGCACTCTTCTGCTGAAGAGACTGATGAACAGTG 4478
 QY 1039 actatgattgcatcaactatltgattgaagagactggggtccgcaagagacttaacagac 1098
 Db 4479 ACTATGATTGCTCAACCACTATTGATGAGAGAGCTGGGGTCCCAAGGACCTTACAGAC 4538
 QY 1099 ataccgctgactggaagagtgctaaacctggtcaactgacgaagacgatagtgtgaa 1158
 Db 4539 ATACCGCTGACTGGAAGAGTGCTAACCTGCTTCACTGACGAGAACAGCTATGTGTGAA 4598
 QY 1159 gtaagagatggtctgagcgagcagtggttgagcggagaccgacagatctggcgagcagc 1218
 Db 4599 GTAAGAGATGAGCTGAGCGAGCGGTGTGAGACGGAGACCCACAGATCTGGCCAGCAGC 4658
 QY 1219 ctgcccgaagagactctgagcgcaaaaggtctgaactcattgacctcagcgaagcttgagg 1278
 Db 4659 CTGCCGGAAGAACTCTGAGCCCAAAAGGCTGAGCTCATGAGCCCTACGCAAGCTTTGCG 4718
 QY 1279 ctgagcgaagggaatccataaacaatlatalacgacagcagatgtgctttgagcgtga 1338
 Db 4719 CTGCGCGAAGGAAATCATATAACATTTATACGACACAGATATGCTTTGGACATGCA 4778
 QY 1339 caagtacaggggacatctalaaacaaaggggttgcttaactcagcagggagggaaata 1398
 Db 4779 CACGTACACGGGGCATCTATAAACAAAGGGGTTGCTTACTCAGCAGGAGGAGAAATA 4838
 QY 1399 aagaacaaaggaatcttaagcctatggaaccccttaacttggcaaaaaggtcgtc 1458
 Db 4839 AAGAACAAAGGAATTTCTAAGCCTATTAAACCTTTTACCTTTGCAAAAAGGCTAGCT 4898
 QY 1459 attatacactgtcctgagacatcaaaagccaaagatccatlatctaaagggaacagatg 1518
 Db 4899 ATTATACACTTCTCTGAGACATCAAGAAAGCCAAATCTCATATTAAGGAAACCAATG 4958
 QY 1519 gctgacgggttgccaagcagggcagccagcgtgttlaacctctgctataatagaagac 1578
 Db 4959 GCTGACCGGTTTGCCAAAGCAAGCAGCCAGGCTGTAACTTTCCTATTAATGAAACG 5018
 QY 1579 cccaaagcccaagaaaccagacagctacacaccttaagaagactggtgcaagagataaaaaag 1638


```

|||||
Db 5019 CCCAAGCCCCAGAACCCAGAGACAGTACACCTTAGAAGACTGCGAAGATATAAAAAAG 5078
QY 1639 atagacaaattctctgagactccggagggagactctataactaatatgaggagaatc 1698
Db 5079 ATAGACCAATTCTCTGAGCTCCGGAGGGACCTGCTATACCTCATATGGAAGGAATATC 5138
QY 1699 ctgcccacaaagaaggttagaataatgtccacagatcatcgtctaacccactaga 1758
Db 5139 CTGCCCCACAAAGAAAGGTTAGATATGTCCAACAGATATCATCTCCACCTACGGA 5198
QY 1759 actaaacacctgacagcagttggtcagaacatccctatcatgltcttgaggctacaga 1818
Db 5199 ACTAAACACCTGCAGCAGATTTGCTCAGAACATCCCTTATCATGTCTGAGCTACAGGA 5258
QY 1819 gtgagctgactggtggtcaaatatgtgtgcctccagctggtttaatgctaactcc 1878
Db 5259 GTGGCTGACTCGGTGGTCAAAACATGTGTGCTCCAGCTGGTTAATCTAATCTCTTC 5318
QY 1879 agaatccctccagaaagagactaaagggaagccacccagcgctcaactggaagtgaac 1938
Db 5319 AGAATACCTCCAGAAAGAGACTAAAGGGAGACCCACCCAGGCGTCTACCTGGGAAGTGAC 5378
QY 1939 ttaactgaggtaaacgcggcttaataacggaacaaatatctatggttttggtagaac 1998
Db 5379 TTCACGTAGAGGTAAAGCGGCGCTTAATACGGAACAAATATCTATTGTTTGTAGACACC 5438
QY 1999 tttagagatggtagagactatctactaagaagaagactcaacggtggtgtag 2058
Db 5439 TTTTCAGATGGGTAGAGCTTATCTCTACTAAGAAAGAGACTTCAACCTGGTGGCTTAA 5498
QY 2059 aaatactgagaaatatttccaaagatttgataatccaaaggtataaggtcagacat 2118
Db 5499 AAATATCTGGAGGAATTTTCCGAGATTGTGAATATCTAAGGTATCGCGTACAGACAT 5558
QY 2119 ggtcagcttgcgttgccaggtlaagtcaggagactggtccagaatatgtggagattg 2178
Db 5559 GGTCACGCTTTGTTGTCGCCAGGTAAATCAGGAGCTGGCCAAATATTTGGGGATTTGAG 5618
QY 2179 aaactgattgcatagaagaccccaaaactcaagacaggtgagaaggttagaagaac 2238
Db 5619 AAAGCTCATTTGTCATACAGACCCCAAGCTCAAGGACAGTGAAGGAGTGAATGAAC 5678
QY 2239 attaaagaaaccttaaccaatlgacacacagagactgagatlaagatgtagtctc 2298
Db 5679 ATTAAGAGAGCCCTTACCAATTGACACAGAGACGCGATTAATGATGTGATCTCTC 5738
QY 2299 ctgacctgtgcttttaagggtgaggaacacccctggaacgtttgggtgacccctat 2358
Db 5739 CTGCCCTTGTGTGCTTTTGTAGGAGACACCCCTGGACAGCTTTGGGCTGACCCCTAT 5798
QY 2359 gaattcttaaggaggacccccccgcttgacagaatgcttgcaagaatgtagtctat 2418
Db 5799 GAATTCCTCTACGGGGGACCCCGCTGTGTAAGAAATGCTTCTGTATATGTCTGTAT 5858
QY 2419 gtgctgttccacacgttctctcttaagctcaagcgctgagtggtgaggaagca 2478
Db 5859 GTGCTGCTTCCAGCGCTTGTCTCTAGGCTCAAGCGCTCGAGTGGGTGAGGACAGA 5918
QY 2479 gegtgaagcagctccgggaagcctacag---agagacttgcgaattccacatgc 2535
Db 5919 GCCTGGAACAGAGCTCCGGGAGGCGTCTACTCAGAGAGAAGGAGACTTCCAAATTCACATCGC 5978
QY 2536 ttccaagtggagatcaatcaatgtatgttagagccacccgagcgggaagaactcgagaatcg 2595
Db 5979 TTCCAAGTGGAGATTCACTCTATGTATTAGCGCCACATGCGAGGAACCTCGAGACTCGG 6038
QY 2596 tgaagaagaaactatctgacttttgacacacacaaagcgctgtaaaatcgaaagaatc 2655
Db 6039 TGGAAAGGGCCTTATCTGACTTTTGTGACACACCAAGCGCTGTGAATTCGAAGGAATC 6098
QY 2656 cccactgtagatcatgcatccacggttaagccggcgcaactcccgatcgagggtgga 2715
|||||

```

```

Db 6099 TCCACTGGATCATGCATCCACCTTAAGCTGGCGCCACCTCCCGACTCGGGGTGAGA 6158
QY 2716 gccgaaagactgaaatcccttaagcttgcctccatcagcgtggtcttaactgtc 2775
Db 6159 GCCGAAGAGCTGAGATCCCTTAAGCTTGGCTTCATCCCTGCTGTTCTTACTCTAAC 6218
QY 2776 aatactccaaatgtaagttaaacgcctgtgtgagacagccggaactccataaacct 2835
Db 6219 AATAACTCCAGCGAGTAGTAAGCGCTTATAGACAGCTGCAAGCCCATATGACCTTT 6278
QY 2836 atctcaactcgttactactgactcctcgtagaaggtatataataataacgaactcaag 2895
Db 6279 ATCCCTTACCTGGCTGATATTGACCTGATACCGGCTGCTACTGTAATAGCACTCGAG 6338
QY 2896 ggaagctcccttgggagactgtgtgcctgaaatataatgtcgtccttcgaatgaatcc 2955
Db 6339 TGTTCCTCTAGAGGACCTGTGTGGCTGAACTGATTTGCTCCGATTTGATTAACCC 6398
QY 2956 tggctcaatgacagagccaccccccagatcgtactcgtgtcttaaggtttacgtttg 3015
Db 6399 CGCTGTTA-----AAAGCACACCTCCCAACCTAGTCCGTAATGAGGTTCTATTGCTG 6452
QY 3016 cccagagcccccaataatgagaalatgtgtgaatccctcaagatctctcttgcagca 3075
Db 6453 CCCAGG---CACAGAGAAAGAAATACTGTGGGGGTTCTGGGAATCTTCTGTAGAG 6509
QY 3076 atgagctgcatcaatcattatgaggaatlggaatgagcagctctccagcaagacg 3135
Db 6510 ATGAGCTGCTTACCTTCAACGATGAGAGACTGGAAATGGCGATCTCTCCAGGACCG 6599
QY 3136 aglaagtactctttgttaa 3156
Db 6570 GGTAAATTTCTCTTTGTCA 6590

RESULT 15
PERY17013
LOCUS PERY17013 7808 bp DNA VRL 14-APR-2000
DEFINITION Porcine endogenous retrovirus gag, pol, env genes.
ACCESSION Y17013
VERSION Y17013.1 GI:3005985
KEYWORDS env gene; gag gene; pol gene.
SOURCE porcine endogenous retrovirus.
ORGANISM porcine endogenous retrovirus.
Virus: Retrovirus: Retroviridae; Mammalian type C
retroviruses: 1-Mammalian type C virus group.
REFERENCE 1 (bases 1 to 7808)
AUTHORS Czauderna, F., Fischer, N., Boller, K., Kurth, R. and Tonjes, R.R.
TITLE Establishment and characterization of molecular clones of porcine
JOURNAL endogenous retrovirus replicating on human cells
JOURNAL J. Virol. 74 (9), 4028-4038 (2000)
MEDLINE 20219394
REFERENCE 2 (bases 1 to 7808)
AUTHORS Tonjes, R.R.
TITLE Direct Submission
JOURNAL Submitted (30-MAR-1998) R.R. Tonjes, Paul-Ehrlich Institut,
FEATURES
Source Location/Qualifiers
1..7808
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
/tissue_type="kidney"
/cell_line="PK15"
260..1834
/gene="gag"
260..1834
/gene="gag"
/codon_start=1
/protein_id="CAI6581.1"
/db_xref="GI:3005985"
/db_xref="SPTREMBL:O73504"
/translation="MCQTVTPTLSTLIDHWTEVRSRAHNLISVQKKGPMQTFCASEMP
TPDVGWSEGTFSNSELIAVKAIIFQTPSSHPDPEYILTWODIAEDPPWPWMLN

```

KPKKPRRIALAGEKNSAEKVEPSRIPETIEEPPEPQVPPPPYPAAGAVG
PSAPGAPVVEGAGRSRGAATPEDEIALIPRTYGPMPGGLOPLQWPS
ADLYNMKTNPESSEDPDLGLVESLMEHJOTDMLQTLQTLTAEREIIEA
RKVPGADRPLOLONEIDMGFLTRGMNYNABEGRESIKYALVIGLRASDR
TNLAKEVWOGPNPSPVFLERLMAFRFPPEZOKSAKVALAPIGASADIR
KILORLEGOAEALROLVRAEKAYYGRETEEEKEDRKEEREEREDROREKULT
KIILAAVEGSSRESDFRKIRSGPSONLGNRLPLDKOCAYCEKEGHWRNCPK
KGNKPRVLALEDDKD
2120 .4726
/gene="pol"
2120 .4726
/gene="pol"
/codon_start=1
/protein_id="CAA76582.1"
/db_xref="GI:3005987"
/db_xref="SPTREMBL:073505"
/translation="MGASISFEQGREPSVNNKPIVLTLOLIDDEYRISPOVKPND
IOSMLQEPQMAETAGMGLAKOVPROVIOUKASATPYSVROYPLSREREGITWPHVO
RLIOGILVPOSPWMTPLPVAKPGSTINDYRPVODLREKNRVDIHPVPPNYLLS
ALPERNMYTVLDKDAFCPLRLHPTSQLFAEMWDPQGTGOLTWRLPOGRNS
PTIFEDALHDLANFRIOHPVTLQYVDLIDAGATKODCEGTKALLLESDIGYR
ASAKKQICREVTLYGSLRGOMRLTAPARKKTVYOIPAPTAPOVRELFATGCR
LMIPTATLAAPLYPLTKEKEFSNAPEHOKAFDAIKKALISAPALAPDMKPTLY
VDERKGVARVLOTGLGPMRRPVAYLSKLDIPASGMPYCLKAIAVALIYDADKLI
LGNITVIAPHALENTYROPDDMMNTNAKMTIOSLLTERVTFAPALNPAATLPE
ETDEPVTHDQHLIETGVKRDLDIPLTEGLVMTFTDSSYVEGKMAAAYVDG
THTIMASSLEPSTSAQKAEMLTALRLABEGSIINTYDSRYAFATAHGAITKOR
GLTISAGREIKNEEILSLLEALHLPLRAIITHCPHOKAKDLISGNOMARVAKOA
AOAVNLPIETPRAPEROYLTEDMOELKKIDOFSEPEGTCYATYCKELLPKES
LEYVOOIHRLHGTGKLOLVRTSPYHLRLPGVADSVKRCVCOULYNAPSRILP
GKRLRSHGAMHEVDFTEVKRAKIKYNTLLVFDVTFSGWEAYPLTKETSTVAKKY
MKRFFRDLXLR
5294 .7267
/gene="env"
5294 .7267
/gene="env"
/codon_start=1
/protein_id="CAA76583.1"
/db_xref="GI:3005988"
/db_xref="SPTREMBL:073506"
/translation="MHPTLSMRHLPTRGCEPKRLRIPLSFASIAFMTLITTPOASSK
RLIDSNPRLPSTLWLIIDPTGVNTSGVAPRGMTWMPHLICRLINAVSTP
PNLVRSYFCGPTGTEKEKYGSGSESPFCRSCVTSNDGDKMPSILODRKFEVN
SGGKRYKMKLYKDKSCSPSDIYLKISTEKGKDNITOKINGSMGIVFKTGGGA
GSTLIRLRIETGTEPPVAVGPDVLADEGPPALPEPMLVPOLTLSPDLTOPSN
GTGLIPLTNPNSPGVPVTKQRLFSLQGAFOAINSDPATSCMTCSGSPPY
EGNAKGGKNVTKEHNRDCTGWSRNKLTLEVSAGCTGCKAPSPHOLCYSTVVEQ
ASENOIYVGNRMWACNTGLTPCVSSSVFNOSKDFCVAVOIVPYYHYHBEVVLEQ
DYRNRPKREPSLTSLAVMLGTAAGTAVGTALITGPOOLEKIGELHAMATDRL
ALEESYNSLESLTSLSEVYLGONRGDILLPIREGLICALEBECCTFYVDHSGAIRDS
MSKLRERLERRRRERADGQMEGPNNSPMWMTLLLSALTLGLVLLLVGPELLIN
KEVAFREVEGAVQIMLVLRQOYOGILSGEITDL"

BASE COUNT 2077 a 1971 c 2047 g 1713 t
ORIGIN

Query Match 86.1%; Score 2857.4; DB 14; Length 7808;
Best Local Similarity 95.3%; Pred. No. 0;
Matches 2992; Conservative 0; Mismatches 136; Indels 13; Gaps 4;

QY 19 acagatgacctctctgacctgagattacaccccaatgagcaaacattttgacctgaa 78
DB 2642 AAGAGATGCTCTTCTTCTGCTGAGATTACACCCCACTAGCCAACTTTTGGCTTCGAA 2701
QY 79 tggagagatccaggtacggaagaccggcagctcaacttgagaccgaactgcccacaag 138
DB 2702 TGGAGAGATCCAGGTACGGGAACACCGGGAGACTCAGCTGAGCCGAGATGCCCCAAGG 2761
QY 139 ttcaagaactccccacacatcttgacgaagccctacacaggaacctggccaactcaag 198
DB 2762 TTCAAGAACTCCCGACCACTTTTGAAGAGCCCTACACAGGAGACTGCGCAACTTCAAG 2821
QY 199 atccaacacctcagtgacctccctccacgtacgtgtagtacctgtctcttgcggaagcc 258

DB 2822 ATCCACACCCCTCAGGTGACCCCTCTCCAGTAGTGATGACCTGCTTGCGGGAGGC 2881
QY 259 accaaacaggaactgcttaagaagtaagaaggaactacgtctggaatttctgaactaagc 318
DB 2882 ACCAAACAGGACTGTTAAGAGTAGAAGGACACTGCTGAATTGTCTACCTAGGC 2941
QY 319 tacaagacctctgctaagaagggcccaagatttcaagaagaggtataacttggggatc 378
DB 2942 TACAGAGCTCTGCTAAGAGGCCCAAGATTTCAGAGAGAGAGTAACTACTTGGGGTAC 3001
QY 379 agtttgcggggcggaagcgaatgctgacgaggaaggaagaaacttgaatccagaata 438
DB 3002 AGTTTGCGGGGGGGAGCAGATGCTGACGAGCAGCAGGAAGAAACTGATCCAGATA 3061
QY 439 ccggcccccaacacagcacaagaagtggagagttttggggacagctggaatttgcaga 498
DB 3062 CCGGCCCAACACACACCAACAAAGTAGAGAGTTTGGGGACAGCTGGATTGTCAGA 3121
QY 499 ctgtgataccggggtttgagacctagcagcccaactctaccgcttaaccaagaagaaa 558
DB 3122 CTGTGATCCCGGGGTTTCCGACCTTAGCAGCCCACTTACCCCTTAACCAAGAAAAA 3181
QY 559 gggggattctctggtgctctgagacacagaaggaattgatacaaaaaaggccctg 618
DB 3182 GGGGATTTCTCTGGGCTCTGAGCAGCAGAGGCAATTGATGATCAAAAAGGCCCTG 3241
QY 619 ctgaagagacccgtctctgacctccctctgaagaaacttaaccttatactatgagat 678
DB 3242 CTGAGGCACTGCTCTGCTGCTCCCTCCCTCACTGAATTAACCTTTATCTATGTGAT 3301
QY 679 gaggcgaaggaagtagcccggaagatttaaacccaacaccttagagacca tggagagacct 738
DB 3302 GAGCGTAAGGAGTAGGCCGAGAGCTTTTAACCCAAACCTAGAGCACTTGAGAGAGACT 3361
QY 739 gttgcctacctgtcaagaagcttgatccctgtagcagtggttgccgatactctgaag 798
DB 3362 GTTGCTACCTGTCAAAGAGGCTTGATCCTTAGCAGCACTGCTGAGCTGCTGAG 3421
QY 799 gctatcgagacgtctgctgacctatggttcaagaagcttgagaaactggaactggaagat 858
DB 3422 GCTATTCGACGCTGTGGCCTACTGCTCAAGACGCTGCAAAATTGACTTGGGACAGAT 3481
QY 859 ataactgtaataagcccccaatgacatgtggaagacatcgctcgagcccccaagacgata 918
DB 3482 ATAACGTGAATAGCCCCCATGATGGAAGAATCGTTCCGAGCCCCCAGACCGATGG 3541
QY 919 atgacccaagcccgatgaaccacatacaagaactgctcttccagaagaggttcaacttc 978
DB 3542 ATGACCAACGCCCGCATGACCCACTATCAAAAGCCGTGTCACAGAGAGGGTCACTTTC 3601
QY 979 gctcaacagccgctctcaacccctgcaactctctgctctgaggaagacatgtaacccagt 1038
DB 3602 GCTCCACCAAGCCGCTCTCAACCCCTGCACTCTTCTGCTGAAGAGACTGATGAACCAAGT 3661
QY 1039 actcaatgattgcatcaactatgattgaaggaagactgggttcgcaaggaacttaccagac 1098
DB 3662 ACTCATGATTGCACTCACTAATTGATTGAGAGACTGGGGGCCCAAGGACTTACAGAC 3721
QY 1099 ataccgctgactggaagaagtgttaacctgttcaacttgacggaagaacagctatgttgaa 1158
DB 3722 ATACCGCTGACTGGAAGAGTCTAACCCTGTTCACTGACGAGAACACTATGTTGTGAA 3781
QY 1159 gtaagagatgagctgagggcgagtggtgagaggaagccgacagatcttgagcagagc 1218
DB 3782 GGTAAAGAGATGGCTGGGGCGGGTGTGTGACACGGGACCAACAGATCTGGGCCACGAC 3841
QY 1219 ctgcccgaaggaacttcaagcgaagaagcttgagcttcaagcctcaagcgaagctttgcg 1278
DB 3842 CTGCGGGAAGGAACTTACGCCCAAAAGGCTGAGCTCATGGCCCTACAGCAAGCTTTGCGG 3901
QY 1279 ctgagcgaaggaatcatataaattatacggagcagagtagtgccttgcagtga 1338

Db 3902 CTGGCCGAGGGAATCCATTAACATTTATACGACAGAGATATGCTTTGGCACTGCA 3961
 QY 1339 cagctacacggggccatctataacaaagggtgtgtctaccccaagcaggaggaata 1398
 Db 3962 CACGTACACGGGGCCATCTATAAACAAAGGGGGTGTGCTTACCTACGAGGAGGGAATA 4021
 QY 1339 aagaacaaaggaatcttaagcctataggaagccttaacatttgcacaaagctagct 1458
 Db 4022 AAGAACAAAGGGAATTTCTAAGCTTATAGAACCTTACATTTGGCAAAAAGGCTAGCT 4081
 QY 1459 attatacgtctcttgacatcagaagaagcgaatctcctatctataggaagcagatg 1518
 Db 4082 ATTATACACTGTCTTGACATCAGAAAGCCAAAGATCTCATATCTAAGGGAACCAAGATG 4141
 QY 1519 gctgacgggtgtgcaagcagcagccagctgttaacctctgtcctcaatagaaag 1578
 Db 4142 GCTGACGGGGTGGCCAAAGAGGAGCCAGGCTGTAACTTCTGCTTATATATAGAAAGC 4201
 QY 1579 cccaagccccagagaccagagcagtagaaccttggcaagagataaaag 1638
 Db 4202 CCCAGACCCCGAGAACCCAGACGACAGTACACCTAGAAAGCTGGCAAGATAAAAAG 4261
 QY 1639 atagaccagtctctgagctccggagggagcctgtcctatccctatctataggaagaaac 1698
 Db 4262 ATGATACAGTCTCTTGAGACTCCGAGGGAGCTGTATACCTCATATGGGAAGGAATC 4321
 QY 1699 ctgcccccaagaagggttagaatatgtcacaagatacatcgtcttaaccacccctagag 1758
 Db 4322 CTGCCCCAACAAGAGGTTAGAAATATGTCCAAACAGATACATGTCTCACCCACCTTAGGA 4381
 QY 1759 actaaacacctgcagcagctgtgtcagacacacccctatcatgtctctgagctaccagga 1818
 Db 4382 ACTAAACACCTGCAGCAGTGTGTAGAACATCCCTTATCATGTCTGAGGCTACCAAGA 4441
 QY 1819 gtgagctacgtgtgtgtcaaacattgtgtccctgcagctgtgttaacttaactccctcc 1878
 Db 4442 GTGGCTACCTCGGTGGTCAACATGTGTGCTGCGAGTGGTAAATGCTTATCCCTTCC 4501
 QY 1879 agaatacctcagaagaagactaaaggagagccacccagcgctacgtgaggaagtgagac 1938
 Db 4502 AGAATACCTCCAGGAAGAAAGAACTAAGGGAAGGCAAGCCACCGGCTACCTGGAAGTGGAC 4561
 QY 1939 ttaactgaggttaagccggtctaaatacggaaacaaatctatctgttttltgtagacac 1998
 Db 4562 TTCACTGAGGTAAAGCCGGCTAATATCGGAAACAAATATCTATGTTTGTGAGACAC 4621
 QY 1999 ttttcagatgggttagaggttactctactaagaagaagagactcaacgtgtgtgctag 2058
 Db 4622 TTTTCAGGATGGGTAGAGGCTTATCTACTAAGAAAGAGACTTCAACCGTGTGGCTAA- 4680
 QY 2059 aaatactgaggaatattttccaagatttggaataccttaaggttaaggttagagcaat 2118
 Db 4681 AAAATACTGAGGAATTTTTCGAGATTTTGGATATCTAAGTATTCGGGTACAGCAAT 4740
 QY 2119 gttccagcttctgtcccaagtgaaatcagggacttggccaagatatttggatattg 2178
 Db 4741 GGTCCAGCTTTTGTGCCCCAGGTAAAGTCAAGGACTGCCAAGATATTGGGATTTGATTGG 4800
 QY 2179 aaactgcattgtgcatagagaccacaaagctcagagcaggttagagaggttagaatacc 2238
 Db 4801 AAACCTGATTGTGATACAGACCCCAAACTCGAGACAGTGAAGAGATGAATAGAAC 4860
 QY 2239 attaaagagacccttaacaaatgacacagagacgtgcatlaatatgtagtgcctc 2298
 Db 4861 ATTAAAGAGACCTTACCAATTTGACACAGAGACTGGCATTAATGATTTGATAGCTCTC 4920
 QY 2299 ctgaccttgtgcttctttaggggttagagaaacacccctggaacgttgggtgacccctat 2358
 Db 4921 CTGACCTTTGTGCTTTTAAAGGTTAGGAACACCCCTGGACAGTTGGGCTGACCCCTAT 4980
 QY 2359 gaattgcttaagggagacccccccgttggcaggaatgcttgcacatagtgat 2418
 Db 4981 GAATTGCTCTACGGGGGACCCCCCGTTGTGTAAGAAATGCTTGTACATAGTGCTGAT 5040

QY 2419 gtgctgtttcccaagccttgttctctagagctcaagcgctcaggtgggttagagcaga 2478
 Db 5041 GTGCTCTTTCCAGGCTCTGTGTTCTTAGGCTCAAGGCCCTGAGTGGGTGAGGACAGA 5100
 QY 2479 gcttgaagcaagctccgggaagcctactcag--aggaagcttgcaggttccacatcgc 2535
 Db 5101 GCGTGAACAGAGCTCCGGGAGGCTACTCAGGGAAGGAGACTTTCAGATTCACATGCG 5160
 QY 2536 ttccaagttggagatcagltcatgtttagagcgacagctgcaggaacctcgaagactcgg 2595
 Db 5161 TTCCAAGTGGAGATTCACTATGTATGTAGACCCACCGCAGGAGAAACCTCGAGACTCG 5220
 QY 2596 tgaagagacattatctcgtactcttgcacacacaaagcgctgtaaaagtcgaagaatc 2655
 Db 5221 TGAAGAGGCCCTTATCTGCTACTTTGACCAACACAGGCTGTGAAATCGAAGGAATC 5280
 QY 2656 cccaactgatalccatgcatalcccaagltlaagccggcgccacactccagattcgggtgga 2715
 Db 5281 TCCACCTGATTCATCATCCACAGTTAAGCTGGGCGCCACCTCCGACTGGGGTGAGA 5340
 QY 2716 gccgaagaagactggaatacccttaagcttgcctccatcgctgttccctactctgc 2775
 Db 5341 GCGAAGAAAGATGAGAAATCCCTTAAGCTTCGCTCCATCGCCTGTTCTTACTCTAAC 5400
 QY 2776 aataactcctcaagttaatggtlaaacgcttgttggacagccgaactccataaacctt 2835
 Db 5401 AATACTCCCGAGCCAGTAGTAAGAGCCTTATTAACAGCTTGAAACCCCATAGACCTTT 5460
 QY 2836 atctcaccctgttacttactgactccggtacaglatlaatatlaacagacccaag 2895
 Db 5461 ATCCCTTACTTGCGTATATTATGACCTGATACGGGTGTCACTGTAAATAGACACGAG 5520
 QY 2896 ggaagctccttgggagacctgtgtgcctgaattatagtctgccttgcataagtaacc 2955
 Db 5521 TGTGTCTCTAGAGGACACCTGGTGGCTGAACTGCATTTCTGCTCCGATTGATTAACC 5580
 QY 2956 tggctcaatgacccagggccacaccccgagtgactcgttgccttgcaggttttaagttg 3015
 Db 5581 CGCTGTTA-----AAGACACACTCCCAACCTTACCTGTAATGTGGGTCTATGCTG 5634
 QY 3016 cccaagaccccccaataatgaagaatattgttgaatcctcaggaatttcccttgcaga 3075
 Db 5635 CCCAGG---CACAGAGAAAGAGAAATACGTGGGGGTTCTGAGGAATCTTCTGTAGAG 5691
 QY 3076 atgagctgcataacttcttaagtatggaatattggaatgtgcagctctcgaagaagac 3135
 Db 5692 ATGGAGCTGCGTCACTCCAAAGATGAGAGACTGGAATGGCCGATCTCTCCAGAGACG 5751
 QY 3136 agtaagttaccttctgttaa 3156
 Db 5752 GGTAAATTTCTCTTTGTCA 5772

Search completed: February 24, 2002, 02:33:44
 Job time: 19935 sec

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	8209	100.0	8709	6	A66553	A66553 Sequence 3
2	8171.4	99.5	8196	6	A66552	A66552 Sequence 2
3	8039.8	97.9	8763	14	PEN1233657	AJ12933657 Porcine e
4	8011.6	97.6	8750	14	PEN133818	AJ133818 Porcine e
5	7965.6	97.0	8918	14	PEN133816	AJ133816 Porcine e
6	7644	93.1	7808	14	PER117013	Y117013 Porcine end
7	7197.2	87.7	8849	14	PEN133817	AJ133817 Porcine e
8	6493.2	79.1	8918	14	PEN293667	AJ293666 Porcine e
9	6282.4	76.5	8132	6	AR130475	AR130475 Sequence
10	6279.2	76.5	8132	4	AF038690	AF038600 Sus scrof
11	6207.8	75.6	8132	4	AF038699	AF038599 Sus scrof
12	5628.8	68.7	7873	6	AX052638	AX052638 Sequence
13	5528.2	67.3	7362	6	AX052634	AX052634 Sequence
14	5489.8	66.9	7333	6	AF038601	AF038601 Sus scrof
15	5489.8	66.9	7333	6	AR130474	AR130474 Sequence
16	5025.8	61.2	6076	6	AX052636	AX052636 Sequence
17	5025.8	61.2	6076	6	AX052647	AX052647 Sequence
18	4986	60.7	8060	6	AR130473	AR130473 Sequence
19	3755.4	45.7	4402	6	AX002804	AX005635 Sequence
20	3376.6	41.1	3482	6	PERENV2	AX002804 Sequence
21	3376.6	41.1	3482	14	PERENV2	Y112239 Porcine end
22	3295.2	40.1	4918	6	AX052637	A66551 Sequence 1
23	2916	35.5	3320	6	A66551	A66551 Sequence 1
24	2548.4	31.0	3025	4	AF147808	AF147808 Sus scrof
25	2416.2	29.4	15845	2	AC079471	AC079471 Mus muscu
26	2295.8	28.0	8655	6	AR136159	AR136159 Sequence
27	2295.8	28.0	8655	14	AF053745	AF053745 Sequence
28	2286.6	27.9	8431	4	AF151794	AF151794 Phascolar
29	2279.6	27.8	8379	14	GU060065	U06065 Gibbon leuk
30	2261.6	27.6	214764	2	AC092712	AC092712 Mus muscu
31	2256.2	27.5	218938	2	AL592224	AL592224 Mus muscu
32	2244.4	27.3	8098	14	PGCGPE	M26927 Gibbon ape
33	2244	27.3	177788	2	AC069310	AC069310 Mus muscu
34	2244	27.3	185872	2	AC068810	AC068810 Mus muscu
35	2237.8	27.3	199883	2	AC067316	AC067316 Mus muscu
36	2236	27.2	207677	2	AL591952	AL591952 Mus muscu
37	2235	27.2	203155	2	AC079473	AC079473 Mus muscu
38	2233.2	27.2	185548	10	AC005743	AC005743 Mus muscu
39	2226.8	27.1	202333	2	AC080144	AC080144 Mus muscu
40	2221.2	27.1	217034	10	AC012147	AC012147 Mus muscu
41	2220.6	27.0	129317	2	AC093330	AC093330 Mus muscu
42	2217.4	27.0	129317	2	AC074337	AC074337 Mus muscu
43	2217.2	27.0	188460	2	AC079560	AC079560 Mus muscu
44	2216.8	27.0	188550	2	AC079547	AC079547 Mus muscu
45	2215.6	27.0	219294	2	AC015797	AC015797 Mus muscu

RESULT	1			
LOCUS	A66553	8205 bp	DNA	PAT
DEFINITION	Sequence 3 from Patent WO9740167.			
ACCESSION	A66553			
VERSION	A66553.1	GI:4538106		
KEYWORDS				
SOURCE	porcine endogenous retrovirus.			
ORGANISM	porcine endogenous retrovirus			
	viruses; Retroid viruses; Retroviridae; Mammalian type C			
	retroviruses; 1-Mammalian type C virus group.			
REFERENCE	1 (bases 1 to 8209)			
AUTHORS	Galbraith,D.N., Haworth,C., Lees,G.M. and Smith,K.T.			
TITLE	PORCINE RETROVIRUS			
JOURNAL	Patent: WO 9740167-A 3 30-OCT-1997;			
	Q ONE BIOTECH LTD (GB)			
FEATURES	Location/Qualifiers			
source	1..8209			
	/organism="porcine endogenous retrovirus"			
	/db_xref="taxon:61673"			

QY 1201 ggggccaattgcagccctccagatctggcccttctctcgcagatctcctataatgga 1260
|||||
Db 1201 gggggcaattgcagccctccagatctggcccttctctcgcagatctcctataatgga 1260
|||||
QY 1261 aaactaacatcccccttctcgcagagatcccccaagcctcaggggttgtagtccc 1320
|||||
Db 1261 AAACTAACCAATCCCCCTTTCGGAGAGATCCCCAACGCCCTCACGGGGTGGTAGAGCCC 1320
|||||
QY 1321 ttatgtctccacagcctctctggatgtgtctcaacagcctctcagacacttca 1380
|||||
Db 1321 TTATGTCTCTCACCAGCCTCTTGGATGATTTGTCAACACCTCTCTCAGACACTCTTCA 1380
|||||
QY 1381 caaccgagagcgaagagaaatctctgttagagcctagaaaaaatgttcctggggccgag 1440
|||||
Db 1381 CAACCGAGAGCGAGAGAGAAATCTGTTAGAGCTAGAAAAAATGTTCTCGGGCCGACG 1440
|||||
QY 1441 gggagccagcgcaattgcataatgtagatctgagatctggaattcccttgactcgcccggt 1500
|||||
Db 1441 GGGCAACCCAGCAATGGAAATGAGATGACATGGGATTTCCCTTACTCGCCCGGTT 1500
|||||
QY 1501 gggactacaacagcgctgaaggttagagagagcctgaatactatcgcagagcctgtgtg 1560
|||||
Db 1501 GGGACTACAAACACGGCTGAAAGTAGGAGAGACTTGAAAATCTATCGCCAGGCTTGGTGG 1560
|||||
QY 1561 cgggtctccggggcgctcaagacgcccactaaattgtgctaaggtlaagagaggtgtg 1620
|||||
Db 1561 CGGCTCTCCGGGGCGCTCAAGACGGCCCACTAATTTGGCTAAGCTAAGAGAGGTGATGC 1620
|||||
QY 1621 agggagccgaagaacccctccctcgatattctctgagaggtcctcagaaagcctcagcggt 1680
|||||
Db 1621 AGGGAACCGAAGAACTCCCTCGGTATTCTTGAGAGGCTCATGGAAGCCTTCAGGGGCT 1680
|||||
QY 1681 tcacccctttgtatccctacccagagcccaagcctcaagtgcgcttgccctcaatgt 1740
|||||
Db 1681 TCACCCCTTTGTATCCTACCTCAAGAGCCCAAGAAAGCCTCAAGTGGCCCTTCATCTG 1740
|||||
QY 1741 ggcagtcggtcctgatatcaggaagaacttcagagactcgaagaggttlaacagaggtc 1800
|||||
Db 1741 GGCAGTGGGCTCTGGATATCAGAAAGAACTTCAGAGACTGGAAGGGTTACAGAGGCTG 1800
|||||
QY 1801 agttacgtgatctagtgagagagcagaagaagtgatctacagaaggtgaacagaaggt 1860
|||||
Db 1801 AGTTACGTGATCTAGTAGAGAGGACAGAGAGGTGATTCAGAGAGGAAACAGAAAGG 1860
|||||
QY 1861 agaagagacagagaagaagaagaagagagaagaagaagaagaagcgtatagaagcgtc 1920
|||||
Db 1861 AGAAGGAACAGAGAAAAGAAAGAGAGAGAGAGAAAGGAGAAAGCGTGATAGAGGCG 1920
|||||
QY 1921 aagaagaacttgaactaagatcttgccgcagtggtgtgaaggtgaagagcagcaggtga 1980
|||||
Db 1921 AAGAGAAATTTGACTAAGATCTTGGCCGAGTGGTTGAAGGGAAGGACAGCAGGAGGA 1980
|||||
QY 1981 gagagagagatcttaagaaaaattagtcagagccctagacagtcagggaaacttggaata 2040
|||||
Db 1981 GAGAGAGAGATTTTGAAGAAATTTAGGTACAGGCCCTTAGACAGTCAGGGAAACCTGGGCAATA 2040
|||||
QY 2041 ggaacccactcgaagaagacagcagtgctgctatgtlaaagaaaaagagcactctggcaaga 2100
|||||
Db 2041 GGACCCCACTCGACAAAGACAGTGTGCTGATTTGAAGAAAAAGGACACTGGGCAAGGA 2100
|||||
QY 2101 actgccccaaagaggaacaaagagacagaggtcctagatcttagaagaagataaagatt 2160
|||||
Db 2101 ACTGCCCCAAGAAAGGAACAAAGACAGAGGTCTTAGCTTAAGAAAGATTAAGATT 2160
|||||
QY 2161 agggagagacgggtctcgagccctcccgagccagcaggttaacttgaaggtgaggggc 2220
|||||
Db 2161 AGGGAGAACGGGTTCCGAGACCCCTCCCGACGCCCAAGGTAACCTTGAAGGTGAGAGGCG 2220
|||||
QY 2221 aaccagttgagttcctgttgatatacggagcagagcaatlaagtgctgtctacaacatag 2280
|||||
Db 2221 AACCAAGTTGATTCCTGTTGATACCGGAGCGGAGCAITTCAGTGTCTCTAACCAATTAAG 2280
|||||
QY 2281 gaaactaagaagaaaaaactcctgtgtgtgtggtgcacagggcaacgctagatcat 2340
|||||

Db 2281 GAAAACTAAAGAAAAAAATCCTGGTGATGGGTCCACAGGGCAACGGCAGTATCCAT 2340
|||||
QY 2341 ggaactccgaagaagaacggttagatctggaggtgggaaggtlaacccactggtttcgtgtca 2400
|||||
Db 2341 GGACTACCCGAAGAACCGTTGACTTGGAGTGGAGGAGGATTAACCCACTGTTCTGTGTGA 2400
|||||
QY 2401 tcccttagtgcagagtagaccctctctagtgtagaacctactgcacagagtgaggagctcaaa 2460
|||||
Db 2401 TCCTTAGTGCAGTAGTACCCTTCTTAGTAGACACTTACTGACCAAGATGGAGCTCAAA 2460
|||||
QY 2461 ttctcttgaacaagaagaaccagaagtgctgtgtgaatacaaaacccactgtgtgtga 2520
|||||
Db 2461 TTTCTTTTGAACAAGAGAACCCAGAGTGTCTGTGAATTAACAAACCCATCACTGTGTGA 2520
|||||
QY 2521 cccctcaattagatgtagatatacgactatcttccccaagttaaagcttgatcaagata 2580
|||||
Db 2521 CCTTCAATTTAGATGATGATATGACATATTCTCCCAAGTAAAGCCTGTATCAAGATA 2580
|||||
QY 2581 tacagtcclgtgtgagcagtttcccaagcctgggcaagaacccgaggtatgggttgg 2640
|||||
Db 2581 TACAGTCTGTGTGAGACACTTCCCAAGCCTGGGCAAAACCGAGGGATGGGTTGG 2640
|||||
QY 2641 caaagaagltcccccacaggttatcaactgaagccagtgctacacagatcaagtica 2700
|||||
Db 2641 CAAGAAGATTCCTCCCAAGAGTTATCACTGAAGGCCAGTGCTACACAGATACGTCA 2700
|||||
QY 2701 gacagtagcccttgtagtagagaggtcgcgaagaagaaatttggccgcatgttcaagaattaa 2760
|||||
Db 2701 GACAGTAGCCCTTGATTAAGAGAGGCTGAGAAAGAAATTTGGCCGCAATGTTCAAAAGTTAA 2760
|||||
QY 2761 tccaacagggcatccctagttccctgtccaaatccctctggaatactccctcgtcaccggtta 2820
|||||
Db 2761 TCCAACAGGGCATCCTAGTTCCTGTCTCAATCCCTTGGAAATACCTCCCTGCTAACGGTTA 2820
|||||
QY 2821 ggaagccttggagccaattgatalcagaccagtagcaagacttgaagaggtlcaataaagg 2880
|||||
Db 2821 GGAAGCCTGGAGCAATGATTTATCGACACAGTACAGACTTGAAGAGGCTCAATTAAGAGG 2880
|||||
QY 2881 tgcaggaatataccccaaggtcccgaaacccctataactctctgtgagccctccgctgt 2940
|||||
Db 2881 TGCAGGACATACACCCCAAGGCTCCCAACCTTATAACCTTGTGAGCGGCCCTCCGCTGG 2940
|||||
QY 2941 aacggaactcgtatacagatattggaacttaaaagatccctctctcgtcgtgagattacac 3000
|||||
Db 2941 AACGGAACCTGCTACACAGATTTGACTTAAAGATGCTCTTCTTGCTCTGAATTTACCC 3000
|||||
QY 3001 ccaactagccaacacttcttgctctcgaattgagagatccaggtlacggaagaacccgggc 3060
|||||
Db 3001 CCACCTAGCCCAACACTTTTTCGCTTGCATTTGAGAGATCCAGGTACGGGAAGAACCGGGC 3060
|||||
QY 3061 agctcaacttggaccgcacttgccccaaaggttcaagaactcccgacacatctttagcgaag 3120
|||||
Db 3061 AGCTCACTGCGGAGCCGACACCTCCCAAGGGGTTCAAGAACTCCCGACCACTTTTGACGAAG 3120
|||||
QY 3121 cccctacagaagagacttggccaacttcaggaatcaaacacccctaagttgacccctcccgat 3180
|||||
Db 3121 CCTTACACAGGAGCTGGGCAACTTGAATCAACACCTTAGGATGACCTCTCCGAT 3180
|||||
QY 3181 acgtgatgaactcgtctcggcgggagccaccaaaagagacttgaaggtlacgaag 3240
|||||
Db 3181 ACGTGATGACCGTCTTCGCGGGAGGACCAACAAAGAGACGCTTTAGAGAGTACGAAG 3240
|||||
QY 3241 caactactgtgaattgtctgaactagcttacagagcctctgtctaagaagcccgatlt 3300
|||||
Db 3241 CACTACTGCTGGAATTTGCTGACCTTAGGCTACAGAGCCTCTCTAAGAAAGGCCAGATT 3300
|||||
QY 3301 ggaaggaaggttaacataacttggggttaagtttcggggcgaggagagatgtgcgaag 3360
|||||
Db 3301 GCAAGGAGAGGTAACATATCTTGGGTACAGTTTCGGGGCGGGAGCGATGTGTCAGGG 3360
|||||
QY 3361 aggcaggaagaagaactgtagtcagatccggcccaacccagccgaacaggttagag 3420
|||||

Db 3361 AGGACGGAGAAACTGTAGTCCAGATACCGGCCCCACACGACCAAAAGTAGAG 3420
QY 3421 agtttttgggaacagcttgatcttgacagcttgatcccggtgttgacaccttaag 3480
Db 3421 AGTTTGGGACAGCTGGATTGTCAGACTGTGATCCGGGGTTGGCAGACTTAGCAG 3480
QY 3481 ccccaactacccgtaacaaagaaagggatctccctggcctccgaagcagaga 3540
Db 3481 CCCCACTCCACCCCTAACCAAAAGAGGGGATCTCTCGGGCTCCCTGAGACACAGA 3540
QY 3541 aggcatttgatctatcaaaaaggccctgtgagcgacactgtctggccctccgaag 3600
Db 3541 AGGCATTGTGATGATCAAAAAGGCCCTGTGAGCGACCTGTGCTCGCCCTCCCTGACG 3600
QY 3601 taactaaaccttaaccttaagtgtatgtagaggttaaggagtagcccgaggagttaa 3660
Db 3601 TAACTAAACCTTTACCTTTATGTGTGATGAGCGTAAGGGGTAGCCCGAGAGATTTTAA 3660
QY 3661 cccaaaccttagaaccatlgagagagacgtgtgactactgtcaaaagaaagctgatactg 3720
Db 3661 CCCAAACCCCTAGACACATGAGAGAGACCTGTGCTACCTACCTGTCAAAGAGCTTGATCTG 3720
QY 3721 tagcaagtgtgtgcccgtatgtctgaagctatcgcaagctgtgcccatactgtgaag 3780
Db 3721 TAGCCAGGTGTGGCCCGTATGTCTGAAGGCTATGCGAGCTGTGGCCATCTGTCTCAAG 3780
QY 3781 agcgtgacaacttgactttgagacagatatgaactgtaataagcccccatgacattgaga 3840
Db 3781 AGCCTGACAAATTGACTTTTGGACAGATATTACTGTATATAGCCCCCATGTGATTTGAGA 3840
QY 3841 acatcgttcggcagccccacagccgaatgagatgacaaagcccgatgacccatcatcaa 3900
Db 3841 ACATCGTTGCGACACCCCAAGACGATGAGATGACCAACGCCCGCATGACCCATATCAAA 3900
QY 3901 gccgtgttctacagagagaggttcaacttgcctccacagccgctctaaacctgtccactc 3960
Db 3901 GCCGTGTTCTCACAGAGGGGTCACTTGTGCTCCACAGCCGCTCTAACCTGCGCCACTC 3960
QY 3961 ttctgacctgaagactatgaacacagtgacatgatttgacataactatttgattag 4020
Db 3961 TTCTGCTGAGAGACTGATGAACCAAGTAGATGATGATGATGATGATGATGATGATGATG 4020
QY 4021 agactgggttcgcaagaagaccttaagaacataccgtgtgactgtgagaagtgtctaaactgt 4080
Db 4021 AGACTGGGTTCGCAAGACACCTTACAGACATACCGCTGACCTGAGAGATGCTTAACCTGCT 4080
QY 4081 tcaactgaagaaagagactatgtgtgaaggttaagagatgagctggggcgagatgtgtg 4140
Db 4081 TCACTGACGAGAGAGACTATGTGTGAGGTAAAGAGATGAGCTGGGGCGGACGTGTGTGG 4140
QY 4141 acggaaccggaacgactgtggccagcagcctgtccggaaggaacttgcgcgaagaagctg 4200
Db 4141 ACGGACCCGCGACGATCTGGGCGCAGACGCTGCGGAAGGAACCTTCAAGCCGCAAAAGCGCTG 4200
QY 4201 agctcatggccctcaagaaacttggcgctgtgcccgaaggaaatccataaacttata 4260
Db 4201 AGCTCATGGCCCTCACGCAACCTTGGCGGTGCGGAAAGGAATTCATTAACCTTTATA 4260
QY 4261 cggagacaggtatgaccttgcagctgacacgtfacagggggcatctataaacaagg 4320
Db 4261 CGGACAGAGGTATGCTTTCGACGTGACACGTCACACGAGGGGCACTTATTAACAAAGGG 4320
QY 4321 ggttgcttacctcagcagggaggaataaagaacaaagaaattctaagcctatag 4380
Db 4321 GGTTGCTTACCTCAGCAGGAGGGAATAAAGAACAAAGAGAAATTCTAAGCTATTAG 4380
QY 4381 aagcctacatttgcacaaaaggttagctatatacactgtcctggacatcagaagaca 4440
Db 4381 AAGCCTTACATTGTCACAAAAGGCTAGATTATATACACTGTCTGGACATCAGAAACCA 4440
QY 4441 aagatcctaatctagaaggaacagatgctgacggttggccaagcaggaaccagg 4500
Db 4441 AAGATCTATATCTAGAGGGAACAGATGCTGACCGGGTTGCCAAAGCAGGACGCCAGG 4500

QY 4501 ctgttaaccttctgacctatagaagaagcccaaaagcccaagaccagacagatgata 4560
Db 4501 CTGTTAACTCTTCTGCTTATATATAGAAACGCCCAAAAGCCCAACACACAGACAGTACA 4560
QY 4561 cccatgaagacttgcagaagataaanaagataagacagtlctctgagactccggaaggga 4620
Db 4561 CCTTGAAGACTGTGCAAGAGATTAATAAAGATAGACAGATGTTCTGTGAGCTCCGAGGGGA 4620
QY 4621 cctgtcataccatataatgggaaggaatccctgcccccaaaaaggttaagatattgtcc 4680
Db 4621 CCTGTATACCTTATATGGGAAGGAATCTCTCCCAACAAAGGGTTAGATATGTTC 4680
QY 4681 aacagaatactgtctcaaccccaactgagaacttaaaacactcagcagatttgcagaacat 4740
Db 4681 AACAGTATATGCTTACACCACTTACCACTTACCACTTACCACTTACCACTTACCACT 4740
QY 4741 cccctatcatgttctgaggttacaagagatgactgagctgggttgatcaaacatgtgtgc 4800
Db 4741 CCCCTATCATGTTCTGAGGGTACACAGAGTGGCTGACTCGGTGGTCAAAACATTTGTGTC 4800
QY 4801 cctgcaagctgttgaatgtcttaactcctccagaatactccagaagaagactaaggaga 4860
Db 4801 CCTGCAAGCTGTTAATGCTAATCTTCCGATATCTCCAGATATCTCCAGAAAGAGACTAAGGGGA 4860
QY 4861 gccaccagcgcctcaccttgcagatgagacttcaactgaggtlaaagccggcctaaatacggaa 4920
Db 4861 GCCACCAAGCGGTCTACCTGGAAAGTGAAGTCACTGATGAGTAAAGCGGCTAATATCGGAA 4920
QY 4921 acaaatatcatattgttcttctgagacacttctcagaatgaggttagaggtctactacta 4980
Db 4921 ACAAAATATCTATGTTGTTTGTAGACACCTTTCAGAGATGGGTAGAGGCTTATCTACTTA 4980
QY 4981 agaaagaagacttcaaacctggtgtgactgaagaataacttgaaggaataatttccaaagattg 5040
Db 4981 AGAAAGAGACTTCAACCCGTTGCTGCTAAGAAATACCTGAGAGAAATTTTCCAAATTTG 5040
QY 5041 gaatacctaaggtatagaggtcagacaaatgctcagcttctgttgcgaagtaagtcagg 5100
Db 5041 GAACTACTTAAGGTAAATAGAGGTGACAGCAATGTCAGCTTCTGTCGCCAGTAAGTACAG 5100
QY 5101 gactggccaagatatgtggagattgtatgtgaactgcatgtgtacacagaccacaagct 5160
Db 5101 GACTGGCCAAAGATATGGGAGATTGATGGAACTGCAATGTCATACAAACCCCAAAAGCT 5160
QY 5161 caggacaggttagagagatgaaatagacacattaaagaaccccttaacaaatgacacag 5220
Db 5161 CAGACAGGTAGAGAGATGAAATAGAACCATTAAGAAAGACCCCTTACCAAAATTTGACACAG 5220
QY 5221 agactggcaatlaagatttgatgtgtcctcctgtctgttcttctttaggttgaggaaca 5280
Db 5221 AGACTGGCAATTAATGATGATGGAGTGTGCTCTCCCTCTTGTGCTTTAGGGTGAAGACA 5280
QY 5281 ccccttgaagaatttggcgtgacccccctatgaatttctcaggggggaaccccccggttgg 5340
Db 5281 CCCCTGACATTAATGATGATGGAGTGTGCTCTCCCTCTTGTGCTTTAGGGTGAAGACA 5340
QY 5341 cagaatatgaccttgcagacatgactgactgacttccagacacttgtctctctagagc 5400
Db 5341 CAGAAATTTGCTTGCACATATAGTGTGATGTGCTGCTTCCAGGCTTGTCTCTAGGC 5400
QY 5401 tcaaggcgtcagatgtgtgagcagcagcgttgaagacagctccgggaagcctactcaag 5460
Db 5401 TCAAGCGCTCGATGGGTGAGACACGAGCTGGAAGCAGCTCCGGGAGGCTCTACTCAG 5460
QY 5461 gaggaagacttgaagttcacatgcttccaaagttggaattcagttcatttgaagacc 5520
Db 5461 GAGGAAGACTTGAAGTTCCACATCGCTTCCAAAGTTGGAATTCAGCTATGTTAGACGC 5520
QY 5521 accgttcagaaacccctcagactcgttgcgaaggaacactatctgtactttgacacac 5580
Db 5521 ACCGTGACAGAAACCTTCGAGACTCGGTGGAAGGAGACCTTATCTGCTACTTTTGTGACACAC 5580

QY 5581 caacgctgtgaaagtcgaaggaatccccactggtacatccatccatcgttaagccg 5640
DB 5581 CAACGGCTGTGAAAGTGAAGGAATCCCACTGGATGTCATGATCCACGTTAAAGCCG 5640
QY 5641 cgcacactccgcactcgggtgagagcgaagaaagactggaatcccccttaagctcgcc 5700
DB 5641 CGCACACTCCGACTCGGGGTGAGAGCCCAAAAAGACTGGAATCCCCCTTAAGCTCGCC 5700
QY 5701 tccatgcctgtgtcttactcttaacataactccccagagccagtaglaagccctata 5760
DB 5701 TCCATGCCCTGTGTTCTTACTCTAACATAACTCCCCAGGCCAGTAGTAAGCCCTTATA 5760
QY 5761 gaagcctcgaagccccatagaacttacccttaactgctgtatattgaacctgaag 5820
DB 5761 GACAGCTCGAAGCCCCCAACCTTATCCCTTACCGGTGATATTGACCCCTGATACG 5820
QY 5821 ggtgtacgtgtaaaactgaactgaggtgtgtctcctagaagcactggtgacctgaag 5880
DB 5821 GGTGTCACTGTAATTAACACTCGAGGTGTGTCTCTAGAGGCACTGTGGCTGAACTG 5880
QY 5881 cattctgcctccgattgaatlaaaccccgctgtlaaaagcacactcccaactagtcgt 5940
DB 5881 CATTTCGCTCCGATTGATTAACCCCGGTGTTAAAGCACACTCCCAACTAGTCGT 5940
QY 5941 agttaaggttctatctgtcgtccagagcacagagaagaatactgtgggtttctgg 6000
DB 5941 AGTTATGGTTCATTGCTCTCCAGGACAGAGAAAGAAATACGTGGGGTTCGTGG 6000
QY 6001 gaatcctctgttaggaatgagagctgctacactccaagatggaacgtggaatggtccg 6060
DB 6001 GAATCCTCTGTAGGAATGAGAGCTGGGTACTCCACATGGAACGTGGAATGGCCG 6060
QY 6061 atctctccagagaccggtlaaacttctcttgcattccggccggcggaagtaaca 6120
DB 6061 ATCTCTCCAGAGCCGGGTAAATTTCTCTTGTCAATTCCGCCGGCGCAAGTACAA 6120
QY 6121 atgatgaactatataaagalaagagctgtcccatcagacttagattatctaaagata 6180
DB 6121 ATGATGAACATATATAAGATTAAGACTGCTCCCATCACTTAATATTCTTAAGATA 6180
QY 6181 agttactgtgaagaaagaaatattcaaaagtgtgataatggtatgagtcggg 6240
DB 6181 AGTTCTGTAAGAAAGAAAGAAATATCAAAAGTGGATTAAGATGAGCGGGG 6240
QY 6241 aatagtttattatataatgagcgggagaggttccacttcaacttgcgcttagata 6300
DB 6241 AATAGTTTATTATATATGCGGGGAGAGAGGTTCCACTTTAACCTTTGGCTTAGATA 6300
QY 6301 gagaacgggagacgaacccccgtgtgcaatgagacccgataaagttactgtgaacaggg 6360
DB 6301 GAGACGGGAGACGAACCCCCGTGTGCAATGGAGACCCGATTAAGTACTGGCTGAACGGG 6360
QY 6361 ccccgccctgtgagccacccgactaactgtgcggtgcccccaatcctcgctgggct 6420
DB 6361 CCCCCGCCCCGTGAGCCACCGCATTAATTGCGGTGCCCAATTAACCTCGCTGGCGCT 6420
QY 6421 gacataacagagccgctgaacagttacactgtgattatccctcaacagccctaga 6480
DB 6421 GACATAACAGAGCCGCTGACCAACATACACGTGATTATTCCTACCAACAGCCCTAGA 6480
QY 6481 aactccccaggtgtctcgttlaagacagagacagactctcaagtcataccagggact 6540
DB 6481 AACTCCCCAGGTGTCTCCTGTTAAGACAGAGACAGACTCTTCAGTCTCAWCCAGGAGCT 6540
QY 6541 ttccaagcatcaactcagaccgactgtgacacttcttctgtgtgttctatcc 6600
DB 6541 TTCCAAGCATCAACTCAGACCCTGATGCACTTCTTGTGTGTGTCTATCC 6600
QY 6601 tcaaggcctcctattatagaaggtgcttaagaagaagaalccaatltgaccaaggg 6660
DB 6601 TCAGGGCCTCCTATTATATAGGGGAGATGCTAAAGAAAGAAATTCATGTGACCAAGAG 6660
QY 6661 catagaatcaatgtacatggtgggtgccgaataagcttaacctcaagtttccggg 6720

DB 6661 CATAGAAATCAATGTACATGAGGGGTCCGAAATAAGCTTACCTCAGTAAGTTCCGG 6720
QY 6721 aagggacatgcataggaagaagctcccccatacccaaacacacttgcataagctgt 6780
DB 6721 AAGGGACATGATAGGAAGAAGCTCCCATCCACCAACACTTGTGATATGATCTGTG 6780
QY 6781 gttatgagcaagccctcagaagaatcagatattagtaactggttataacaggtgtgca 6840
DB 6781 GTTATGAGCAGGCTCCACAAATATCAGTATTAGTACCTGTATTATAACAGGTGGGA 6840
QY 6841 tgcataactgggttlaacccccgtgtttccacctagcttccaaacatccaagaattc 6900
DB 6841 TGCATACTGGGTAAACCCCTGTGTTCACCTCAGCTTCAACCAATCCAAAGATTTC 6900
QY 6901 tgtgtcatgttccaatgttccccgagtgatcactccactcgtgaagaggtgtctgtat 6960
DB 6901 TGTGTATGTGTCAAATCTCCCGAGGTATCTACATCTGAGGAAGTGTCTTAT 6960
QY 6961 gaatatgactatcgtlatataccgacaaagagaacccgtatcccttaacctagctga 7020
DB 6961 GATATGACTATCGGTATACGACCAAAAGAGACCCGTATCCCTTACCTAGCTGTA 7020
QY 7021 atgtctgattttaggaacggcggtgtggttaggaacaggaacgtgtccctgatacaaga 7080
DB 7021 ATGCTCGATTTAGGACCGCGCTGGCGTAGAAGAGGACAGTGCCTGATCACAGA 7080
QY 7081 ccaagcagctagaagaagagactgtgtgactacactgtgcccatacagaagaatccga 7140
DB 7081 CCACAGCACTAGAAAGAGACTTGTGTAGCTACATGCGGCATACAGAAAGATCTCGA 7140
QY 7141 gccctlaagagctcgttagcaacctagaagaagtcctgaacttctgtcgaagtgt 7200
DB 7141 GCTTAAAGAGACTGTTAGACCACTAGAAGAGTCCCTGACTTCTTGTGAGTGTG 7200
QY 7201 ctacagaacccggagggagttagatctgttcttcaagaagaaggtgtgtatgtgacgc 7260
DB 7201 CTACAGAACCCGGAGGGGATTAGATCTGCTGTTTAAAGAAAGTGTATGTGACACC 7260
QY 7261 ttaagaagaagaatgtgtctctatgtatagatacactcaagagccatcaagactccaag 7320
DB 7261 TTAAAGAAAGAAATGTTGCTTATGTATGATATCACTAGAGCCATCAGAGACTCATG 7320
QY 7321 aagcttaaaaaaagttagagagcgtcgaagggagaagaagagctgacacgaaggtgt 7380
DB 7321 AAGCTTAAAAAAGTTAGAGAGGCGTGAAGGCAAGAGGCTGACACAGGGGTGTT 7380
QY 7381 gaaggaatgttcaacaggtctccttga tgaaccaactgtcttctgtctgaacggggcc 7440
DB 7381 GAAAGATGTTCAACAGGTCCTCTTGATGACACACCTGCTTGTGCTGACGGGGCC 7440
QY 7441 ctgagtgtcctgtcctgttacttaagttgagctgtgaccttaataagttgtgtgc 7500
DB 7441 CTGAGTGTCTGCTCTCTTACTTAACTAGTGGGCTTCTTAATTAAGTGTGTGTC 7500
QY 7501 ttgtttagaagacaggtgagtgcaatccagatcaatgtaagtaagcaagttaccaagc 7560
DB 7501 TTTGTTAAGAAAGAGTGAAGTCAATGATGATGATGATGATGATGATGATGATGATG 7560
QY 7561 ctctgagccaaggaagaaactgaaccttagcttcccaagttctaagattagaactataa 7620
DB 7561 CTCTGAGCCCAAGGAAGAACTGACCTTAGGCTTCCAGTCTTAAGATTAGAACTATTA 7620
QY 7621 caagcaagaaggtgaggaatgaagaagatgaagaatgaacacttaacccctccgaacccag 7680
DB 7621 CAAGCAAGAAGTGGGATGAAGAGATGAAGAAATGCAACTTAACCTCCAGAACCCAGG 7680
QY 7681 aagttaatataaagctcctaaatgcctccgaatctcaagacccgtgctgctgcaataa 7740
DB 7681 AAGTTAATAAAGCTCTAAATGCCCGGATTTACAGACCTGCTGCTGCTGCTGCTGCTG 7740
QY 7741 gtagaaggtcacacttctatgttccaaggtgtctatctgtgcttaagtaagataac 7800

```

Db 7741 GGTGAAGTACACACTTCTATTGTTCCAGGGCTGCTATCTGGCCTAAGTAATAC 7800
Qy 7801 aggaatagtgtaactaatcgtatctgattctgttaaaactgagcaacatagaag 7860
Db 7801 AGGAATAGTGTGACTAATCGCTAATCTGATCTGTAAACTGACTGGACCATAGAAG 7860
Qy 7861 aattgatacaattgacagccctagtgtactatctcaactgaaatctgtactctgcc 7920
Db 7861 AATTGATTACACATGACACCCCTAGACCTATCTCAACTGCAATCTGTCACTCTGCC 7920
Qy 7921 aggaagccagcagatccgagctccgagcctaatttaaatgattgtccacagagcgc 7980
Db 7921 AGGAGCCAGCAGATCCGAGACTCCGAGACTTTTAAATATTGTCTCAGAGAGCCG 7980
Qy 7981 gggctctgataatttaaaatgattgtccacagagcgcgcctctgatalttaaat 8040
Db 7981 GGGCTCTGATTTTAAATGATGGTCCATGAGCGCGGCTCTGATATTATTAAAT 8040
Qy 8041 gattgtgttgagcagcagcttctgtgtgaaccccaataaagctgtccagatccgca 8100
Db 8041 GATTGTGTGTGAGCAGCAGGCTTTGTGTGAACCCCATTAAGCTGCTCCGATTCGCA 8100
Qy 8101 ctgagagccagctcctcctcctccctcgtgtgtgtaagactgtggcccgagcgtcttg 8160
Db 8101 CTGAGGCGCCGAGTCTCTACCTCCCTGCTGTGTGAGACTGTGGCCCGAGCGCTTGG 8160
Qy 8161 aataaataccctcctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 8209
Db 8161 AATAAATAATCCTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8209

RESULT 2
LOCUS A66552 8196 bp DNA PAT 29-MAR-1999
DEFINITION Sequence 2 from Patent WO9740167.
ACCESSION A66552
VERSION A66552.1 GI:4538105
KEYWORDS
SOURCE porcine endogenous retrovirus.
ORGANISM porcine endogenous retrovirus.
VIRUSES; Retroviral viruses; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
REFERENCE 1 (bases 1 to 8196)
AUTHORS Galbraith,D.N., Haworth,C., Lees,G.M. and Smith,K.T.
TITLE PORCINE RETROVIRUS
JOURNAL Patent: WO 9740167-A 2 30-OCT-1997;
FEATURES
source Q ONE BIOTEC LTD (GB)
location/Qualifiers
1..8196
/organism="porcine endogenous retrovirus"
/Db_xref="taxon:61673"
BASE COUNT 2165 a 2061 c 2146 g 1820 t 4 others
ORIGIN

```

Query Match 99.5%; Score 8171.4; DB 6; Length 8196;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 8191; Conservative 4; Mismatches 0; Indels 2; Gaps 2;

```

Db 181 GACTTGAGAGTAAGGATGCCCTTTGGAACGTGTGTGTGTGTGTGTGTGTGTGTGT 240
Qy 253 ctgagtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 312
Db 241 CTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
Qy 313 ggaactggaagactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 372
Db 301 GGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
Qy 373 agagcccgagagagtgagagagagagagagagagagagagagagagagagagagagag 432
Db 361 AGGCCCCGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
Qy 433 accgagttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 492
Db 421 ACCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
Qy 493 gctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 552
Db 481 GCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
Qy 553 tcttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 612
Db 541 TCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
Qy 613 ttagttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 672
Db 601 TTAGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
Qy 673 aggttaagaagagagagagagagagagagagagagagagagagagagagagagag 732
Db 661 AGGTTAAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Qy 733 gatgtgcatacagagagagagagagagagagagagagagagagagagagagagagag 792
Db 721 GATGTGCATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Qy 793 ttcagactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 852
Db 781 TTCAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
Qy 853 tggcagaagatcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 912
Db 841 TGGCAGAAGATCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 900
Qy 913 cccgaatcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 972
Db 901 CCCGAATCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
Qy 973 cctgtatctaccccgagatcgaagagagagagagagagagagagagagagagagag 1032
Db 961 CCTGTATCTACCCCGAATGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Qy 1033 ccaccccttaccagagagagagagagagagagagagagagagagagagagagagag 1092
Db 1021 CCACCCCTTATCCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Qy 1093 tggtagagagagagagagagagagagagagagagagagagagagagagagagagag 1152
Db 1081 TGGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Qy 1153 acgagatcgagatattacagctgtgcacatattgtccctccagagagagagagagag 1212
Db 1141 ACGAGATCGGATATTACAGCTGTGCACATTTAGGAGAGAGAGAGAGAGAGAGAGAG 1200
Qy 1213 agccctcagagagagagagagagagagagagagagagagagagagagagagagag 1272
Db 1201 AGCCCTCAGAGATTTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
Qy 1273 ccccttctcagagagagagagagagagagagagagagagagagagagagagagagag 1332
Db 1261 CCCCTTCTCGAGAGATGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320

```

QY	1333	acgcgcctacttggatgatgttgcacagcctgcctgcagacactcttcaaacgcgaagac	1392
Db	1321	ACCGACCTACTTGGGATGATGTTCACACACTCTGTAGACACTCTTCCAAACCCAGGAGGC	1380
QY	1333	gagagaaattctgttagaggtcagaanaaaatgttcctggyggccgaggggagccaccgc	1452
Db	1381	GAGGAGAAATTCGTATGAGAGCTAGAAAAAATGTTCTGGGGCCAGCGGGAGCCACGC	1440
QY	1453	agttgcnaaatgatgatgaatbgyatcttcctctgactgcgcccggtttggaactacaac	1512
Db	1441	AGTTGCAAAATGACATTGACATGGGATTTCCCTTGACTGTCGCCCGCTTGGGACTACACA	1500
QY	1513	cgctctgaagtlagggagagcttgaaaatctatgcgcgaagcttctgtgcggttccggt	1572
Db	1501	CGGCTGAAGGAGGAGAGAGCTTGAATAATCTATCGCAGAGCTGTGTTGGCGGCTCCGGG	1560
QY	1573	gcgcctcaagacggtccacctaatttgcctcaagtlagagaggtgaatgcagggaccgaa	1632
Db	1561	GGCGCTCAAGACGGCCCACTAATTTTGGCTAAAGAGAGGTGATGACGAGGAGCCGAACG	1620
QY	1633	aacctccctgcgtatcttcttgaagagccatbgaagcccttgaagcgttccaccctttg	1692
Db	1621	AACCTCCCTCGGATTTTCTTGAAGAGCTATATGAAACCTTCAAGCGGGTTCAACCTTTTG	1680
QY	1693	alccctacctcagaggtccccagaaaagcctcagcttgcccttgccttcatctggcagtc	1752
Db	1681	ATCTCACTCAGAGGCCCAAGAACCTCAGTGGCCCTGGCTTCAATTGGCGAGTGGGCTC	1740
QY	1733	ttgatalcaggaagaacatctcagagactbgaaggtttacagaagcttgaatgaatgac	1812
Db	1741	TGGATATCAGGAAGAAACTTTCAGAGACTGGAAGGGTTACAGAGGCTCAGTTACGTGATC	1800
QY	1813	tafagaaagagggcaggaagaagtgttatcaagaaggagacaaagaagagaagaacaga	1872
Db	1801	TATGTAAGAGAGGCGAGAAAGGTTATTACAGAAAGGAGACACAAAGAGAGAGAGACAGA	1860
QY	1873	gaaaagaagaaggaagagaagaagaagaagaagaacgtbatalgaacgcgaagaagaat	1932
Db	1861	GAAGAAGAAAGGAGAGAGAAAGGAGAAAGGAGAGAAAGAGCTATGACGCGCAAGAAAT	1920
QY	1933	tgactaagatcttgcgcgaagtgttctgaagggaagaacagcagaggaagagaagagat	1992
Db	1921	TGACTAAGATCTTGGCCCGAGTGGTTGAAGGGAGAGCGACGAGGAGAGAGAGAT	1980
QY	1993	ttagaanaattagtgagcggccctcagaagcagtcaaggaaccttggcgaatagagcccaacg	2052
Db	1981	TTAGCAAAATTAGGTCTAGCGCCCTTAGACAGTCAAGCAACTGGGCATATGAGACCCACTCG	2040
QY	2053	acaagagaccagtgcgtatcttgaagaaagaacacgtggcgaagagacttcccaca	2112
Db	2041	ACAAGGACCACTGTGCTATTGTAAGSAAAMAAAGACACTGGGCAAGAACTGCCCAAGA	2100
QY	2113	agggaaacaagaagcaggaaggtctcagctctagaagaagaataaagtlagggagaaggg	2172
Db	2101	AGGGAAACAAGGAGACCCAGA-GTCTAGTCTCTAAGAAAMATTAAGGAGCAAC-CG	2158
QY	2173	gttgcgaacccctcccggaagccgaaggtiaacttgaagtlgaaggggaacacagttgt	2232
Db	2159	GTTGCGACCCCTCCCGAGCCCAAGGTTAATTGTAAGGTGAGGGGCAACCTAGTTGAT	2218
QY	2233	tcctgtgtgatccggagcggagcatctcagctgtctgtacaaacatlaggaanaactaaag	2292
Db	2219	TCCTGTGTGATACCGGAGCGGAGCATTCATGCTGTCTACACCATATTAGSAAACATAAAG	2278
QY	2293	aaaaaaatctctgggtgatbgygttccaagggcaacggaatatacgaatgactaccga	2352
Db	2279	AAAAAAATCTCTGGGTATGGGTGCCACAGGCGCAAGCGAGATATCCATGACTATCCGCA	2338
QY	2353	gaacgtgtgacttggaggttggagcgggttaaccacactcgtttctgtfcaatccttagtgc	2412
Db	2339	GAACGTTGACTCTTGAGGTGGAGGGGTAAACCACTGTTTCTGGTCATCCTTGATGTC	2398

QY	2413	cagtaaccctcttagtgaagactactctgaccaagatggggtctaaattcttttgaac	2472
Db	2399	CAGTACCCCTTCCTCAGGTAGAGACTTACTGACCAAGATGGAGCTCTCAATTTCTTTGAAc	2458
QY	2473	aaggaagaaccaagatctcgtgtaatcaacaaccctaaactctgtgttgaccctcaatga	2532
Db	2459	AAGGAAGACCAAGAAgTGTCTGTAAATACAAACCCATCCTGTGTGACCTCCAAATTAG	2518
QY	2533	atgtatgaatlcagactaattctcccaagtaagactctatcaagataaacgtctcgt	2592
Db	2519	ATGTATGAATATACGACTATATTCTCCCAAGTAAAGCTGTATCAAGATTTACAGTCTGT	2578
QY	2593	tggagcagtttccccaagcctgggcsgaaacgcgaaggaatgggtttgcaaacgaatgc	2652
Db	2579	TGGAGCAGTTTCCCAAGCCTTGGGCAGAAACCGCAGAGATGGGTTTGGCAAAAGCAATTC	2638
QY	2653	ccccacaggtattcaactctgaagcgcaagtcacacagatatacgttcaagacagtcctc	2712
Db	2639	CCCCACAGGTATTCAACTGAAGGCCAGTGCTACACCAATATCAAGTACAGCAATACCCCT	2698
QY	2713	tgaagtagaagcctcsgaagaagatcttgccgcatcttcaaaagatlaatccaacagggca	2772
Db	2699	TGAGTAGAAGAGCCTCGAAGAGAAATTTGGCGCATTCCTCAAAAGATTAAATCCAAACAGGGCA	2758
QY	2773	tcctagtctcgttccaaatcccccttgtaatactccccctgctacccgttgaagacgttga	2832
Db	2759	TCCTAGTTCCTGCTCCAAATCCCTCTTGGAAATCTCCCTGCTACCGGTATGGAAAGCCTTGGGa	2818
QY	2833	ccaatgatatcgcgaccgaatacaggacttgaagagatgaaataaaaggttgcaaggaataac	2892
Db	2819	CCAAATGATTATCGACACAGTACAGACTTGAGAAAGGTCAATAAAGGGTGCGAGCAATAC	2878
QY	2893	accacaagctccgaacccctataacctcttgagcgccctccgcgcctgaacggaacttgc	2952
Db	2879	ACCCACAGGTCCTCCCAACCTTATTAACCTCTTGAAGCCCTCCCGCTCAACGGAAGTCTGT	2938
QY	2953	acacagtatctgacttaaaagatgccttctcctcgtsgatatacaccaccactagcaac	3012
Db	2939	ACACAGTATTGAGACTTAAAGAGTGCCTTCTTCCGCTGAGATTATACCCCACTACCAAC	2998
QY	3013	caactttgctcttgaaatgtagagatctccagtaacggaagacgggcgcactctctga	3072
Db	2999	CACTTTGTGCTTGGAAATGAGAGATCCAGTACGGGAAGAACCGGCAAGCTCACCTGGA	3058
QY	3073	cccgactcccccaagggttcaagaactccccgcacactcttgaagaaagccctacaagaag	3132
Db	3059	CCCGACTCCCCCAAGGGTTCAAGAACTCCCCGACACTCTTTGACGAAGGCCCTACACAAAGG	3118
QY	3133	acctggccaacttcaagatacccaacacctcaagtgaacctcctccagtaactgtatgaac	3192
Db	3119	ACCTGGCCAACCTTCAGGATCCCAACACCCTCAGGTGACCTCTCCACGTACGTGGATGAC	3178
QY	3193	tgcctctggcgggagcccaaaccaagactgccttaagaagtacgaaggaactacactgcgg	3252
Db	3179	TGCTTCTGGCGGGAGCCAAACAGGACTGCTTAAAGGTATCAAGAGCATCACTGCTGGG	3238
QY	3253	aattctcgaactaggtctacagagccctctgtaagaagggccagatcttcaggaagagag	3312
Db	3239	AATTGTCTGACTTAAGCCTACAGAGACCTCTGCTAAGAAAGGCCCAAGATTTCAGAGAAAGG	3298
QY	3313	taacataacttggggatcacagtttgcggggcgggcagcgatgctgacggaagcaacgaaga	3372
Db	3299	TAACTATCTTGGGGTACAGTTTGGCGGGCGGAGCGATGGCTGACGGAGCCACGGAAGA	3358
QY	3373	aaactgtatgctcagataccggcccccaaccacagcccaaaacgaatgagaaggttttgggga	3433
Db	3359	AAACGTATGTCACATACCGGCCCAACACACGCAAAACAAGTGAAGAGATTTTGGGGGA	3418
QY	3433	gaagctggatcttgcaagacttggatctccgggggttgcgaacttaagcagcccaacttacc	3492
Db	3419	CAGCTGGATTTTGCAGACTGTGGATCCCGGGGTTTGTGCAGCTTAAAGACGCCCACTCTAC	3478
QY	3493	cgcttaaccaagaanaaagggtatctctcctgggctccctgaacccaagaaggtatttga	3552

Db 3479 GCGTACCAAGAAAAAGGGGATTCTCTGGGCTCCGAGACACAGAGCATTTGATG 3538
Qy 3553 ctatcaaaaagccctgtgagcgcaactgtctctgcccctccctgacgttaactaacct 3612
Db 3559 CTATCAAAAAGCCCTGCTGAGCCGACACTGCTCTGGCCCTCTCCGACGTAACTAACCT 3598
Qy 3613 ttacccttatgtgatagacgtaagagtagcccgagaggttlaaccnaaacctag 3672
Db 3599 TTACCCCTTATGTGGATGAGCGGTAAAGGAGTAGCCCGAGGAGTTTAAACCAACCCGTAG 3658
Qy 3673 gacatagagagagacactgtgttccactgttcaagaagactgtatccgttagccagtgt 3732
Db 3659 GACCATGAGAGAGACCTGTCTCCCTACCTGTCATAAGAAAGCTTGTATCCGACGTGT 3718
Qy 3733 ggcacctatgtctgaagactatcgcaactgtgtgacatctgtgtaagagacgtacaat 3792
Db 3719 GGCCCGTATGTCTGAGAGCTTCCGACGCTGTGGCCATACGTGTAAGAGACGCTTACAAAT 3778
Qy 3793 tgaacttggagacagaataaactgttaatagcccccaatgcatlgtgagaacatcgttcgc 3852
Db 3779 TGACTTTGGAGACAGATTAATTAATAGCCCCCATGCAATTGGAGAACATCGTTGGCC 3838
Qy 3853 agcccccagacagatgataagcaagcccgatagaccactatcaaacgctgtctca 3912
Db 3839 AGCCCCAGACCGATGATGACCAAGCCCGCATGACCCACTATCAAAAGCCTGCTTCTCA 3898
Qy 3913 cagaagaggtcaacttcgctccacacagccgctcaacccctgcaacccctctgctaaag 3972
Db 3899 CAGAGAGGCTACCTTGGCTCCACACGCGCTCCAAACCTGCCACCTCTTCTGCTGAAG 3958
Qy 3973 agactgtatgaacagactgaactatgatactgcaactatgatatgagagagactgagttcc 4032
Db 3959 AGACTGTATGAACAGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4018
Qy 4033 gcaagagccttaacagacataccgtgtgctgagagagtgctaaactgttcaactgaagaa 4092
Db 4019 GCAAGGACCTTACAGACATCCGCTGACTGAGAGAGTGTACCTGTTTACACGACGAA 4078
Qy 4093 gcaactatgtgtgagagtaagagatgctgagggcgagtggtgagagagagagagagag 4152
Db 4079 GCACCTATGTGTGTAAGTAAGAGATGCTGCTGGGGGCGAGTGGTGAGCGGAGCCGCA 4138
Qy 4153 cgaatcggccagcagctgcccgaagagaaacttaagcgcaaaagctgagctcatgccc 4212
Db 4139 CGATCTGGGCGCAGCGCTGCCGGAAGAGACTTCAGCGCAAAAGGCTGACTCATGGCC 4198
Qy 4213 tcaagcagacttgcgctgagccgaagagaaatccataaacaatltatagagacagagt 4272
Db 4199 TCACGCAAGCTTTCGGCTGGCCGGAAGGAATCCATTAACATTTATACGACACGAGGT 4258
Qy 4273 atgacctgtgactgacagatcaacgagggcatcttaaaagaaggggttgcctacat 4332
Db 4259 ATGCTTTGGGACTGCGACAGCTACCGGGCCATCTATAACAAAGGGGTTCTTAACT 4318
Qy 4333 caagcagggaggaataaagaacaaagagaaatcttaagcctatagaacccatlaact 4392
Db 4319 CAGAGGAGGAGGAATTAAGAAACAAAGGAATTCATTAACCTATTAGAAAGCCTTACAT 4378
Qy 4393 tggccaaaaggtagctatatactgtctgtgacatcaagaagccaaagatcctaat 4452
Db 4379 TGGCAAAAAGGCTAGCTATTATACACTGCTCTGACATCAAGAAAGGATCTCATAT 4438
Qy 4453 ctagaaggaacacagatgctgacccgggttgcgaagcagccagcagctgttaaccttc 4512
Db 4439 CTAGAGGAGACAGATGCGCTGACCGGGTTCACAGACGCGACGCCGCTTTAACTTC 4498
Qy 4513 tgcctataatagaacagcccaagcccaagacccagacagacagtaacacactagaagact 4572
Db 4499 TGGCTAATATAGAAAGCCCAAGGCCCAAGACCCAGAGACGATGACCTTGAAGAGCT 4558
Qy 4573 ggcagaagataaaaaagataagacagttctctgaagctccgagggagacctgtataact 4632
Db 4559 GGCAGAGATTAAGAAATAGACCATGCTCTGACACTCCGAGGAGGACCTGCTATACCT 4618
Qy 4633 catatggaaagaaatcctgcccccaagaaggttaaatatgtccaaacagatataac 4692
Db 4619 CATATGGAGAGAAATCTGCCCCACAAAGAGGTTAATAATGTCTCAACAGATACATC 4678
Qy 4693 gtcatacccaactagaacataacacactgcaagagttgttgcagaacatccctatcatg 4752
Db 4679 GTCTAACCCACCTAGAGACATAAACACCTGACAGAGTGGTTCAGAACATCCCTTATCATG 4738
Qy 4753 ttctgaggtacccaagagatgctgactggtgtgtcaaacatgtgtgtccagctgtg 4812
Db 4739 TTCTGAGGCTACCCAGAGAGTGGCTGACTCGTGTCTAAACATTTGTGCTCCAGCTGG 4798
Qy 4813 ttaatgtcaatcccttcagaataaccccaaggaagagactaaagggaagcccaagcg 4872
Db 4799 TTATGCTATCTCTTCAGATATCTCCAGAAAGAGACTTAAGGGAGAACCCACAGGCG 4858
Qy 4873 ctcaactgggaagtgagacttcaactgagtaagccgctaaatagcgaacataatctat 4932
Db 4859 CTCACGAGGAGAGTGCATTCACAGAGTAAACCCGCTAAATACGAAACAAATATCTAT 4918
Qy 4933 tggtttttgaagacacttltcaagatggtgtagaggttactactaagaagagact 4992
Db 4919 TGGTTTTTGTAGACACTTTTCAAGATGGGTAGAGGCTTATCTTAAGAAAGAGACTT 4978
Qy 4993 caacgctgtgtgtaagaaatactgagagaaatttccaagatttgaataactaagg 5052
Db 4979 CAACCTGGTGGCTTAAGAAATATCTGGAAGAAATTTTCCAAAGATTTGAAATCTTAAG 5038
Qy 5053 taataggtcaagacaatgttccagcttctgttgcacagtaagtaaggaactgccaaga 5112
Db 5039 TATATAGGTCAGACATGCTCCAGCTTTCGTTGCTGAGTAAAGAGGACCTGGCAAGA 5098
Qy 5113 tattggagattgaattggaactgcatltgtcaatagaccccaagaactcagaagtag 5172
Db 5099 TATTGGGATTTGATTTGAAATGCAATTTGTGCTACAGACCCCAAGAGCTCAGACAGTAG 5158
Qy 5173 agagagtagaataagacatlaaagaaccccttaacaaattgaaccaagaagctgcat 5232
Db 5159 AGAGGATGATGAAACATTAAGAGACCTTAACCAAAATGACCAAGAGATGGCATTA 5218
Qy 5233 atgattgtatgtctccctgcttgccttgccttgccttgccttgccttgccttgccttgc 5292
Db 5219 ATGATTGATGATGCTCCTCCGCTGCTTGTGCTTTTAAAGGTAGAGAACCCCTGGACAGT 5278
Qy 5293 ttggagctgaccccatgaattgctctcaggggagcccccccgcttgcagaaatgtcct 5352
Db 5279 TTGGGCTGACCCCTATGAAATGCTCTACGGGGAGCCCCCCTTGGCAGAAATTTGCT 5338
Qy 5353 ttgcacatagtgctgattgtgctgttccagccttgccttgccttgccttgccttgccttgc 5412
Db 5339 TTGCACATATGCTATGCTGCTGCTTTCACACCTTTTGTCTAGGCTCAAGGGGCTCG 5398
Qy 5413 agtgggtgagcagcagcgttgaagcagctcccgagagccctactcaagagagagactg 5472
Db 5399 AGTGGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5458
Qy 5473 aagttccacatcgtctcaagttgagatltcaagttatgttaagcagccagtgagaa 5532
Db 5459 AAGTTCCACATCGCTTCCAAAGTTGAGATGCTATGTTAGAGCGCCACCGCTCAGAA 5518
Qy 5533 acctggaactcgtgtgaagagagacttctcgtactttgaaccaaccaagcgtgtga 5592
Db 5519 ACTGGAAGCTCGGTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5578
Qy 5593 aagtcgaaggaatcccaacccctgagatcaatgcaacagtttaagcagcgccagcactccg 5652
Db 5579 AAGTGAAGGAATCCCCACGCTGAGATCAATGATCCAGATTAACGCGGCGCAGCTCCG 5638
Qy 5653 actcgggttgaagagcgaanaagactggaatcccttaagcttcgctcactcagctgtg 5712
Db 5639 ACTCGGGGTGAGAGCCGAAAGAGTCCCTTAAGCTTCGCTCATGCTGCTGCTGCTGCTG 5698

QY 5713 ttcctactctaaacaaatacctccagccagtaglaaacgccttatagacagctcgac 5772
|||||
Db 5699 TTCCTACTCTAACAAATACCTCCAGGCCAGTAGTAACCCCTTATAGACAGCTCGAAC 5758
QY 5773 ccccaataagccttataccttaacctgctgattatagcccttgatagcgggtgtaactgta 5832
|||||
Db 5759 CCCCATAGACCTTTATCCCTTACCTGCTGATTTAGACCCCTGATACCGGGGTGCTACTGTA 5818
QY 5833 aatagactcgaagtggtgctccttaagagcaaccttgagccttgaaactcattctgctc 5892
|||||
Db 5819 AATAGACTCTGAGGTGTTGCTCTAGAGGCACTGCTGCTGAACTGCATTTCTGCTC 5878
QY 5893 cgattgataaacccgcgtgtlaaaagcaacctcccaacctagtcgtagtattggtgctc 5952
|||||
Db 5879 CCATTGATTAACCCCGCTGTATARAGCACACCTCCCAACCTAGTCCGAGTATAGGCTTC 5938
QY 5953 tattgctcccaagcagcagagaagaataactgtgtgggttctggggaaactcctctgt 6012
|||||
Db 5939 TATTGCTGCCAGGACACAGAAAGAAATACCTGGGGGCTTCTGGGGAAATCCTTCTGT 5998
QY 6013 aagagatagagctgctgaacctccaaagcagtagagactlgaaatlgcgcactcctccag 6072
|||||
Db 5999 AAGAGATGAGCTGCTGCTCCTCCACGATGAGACTGGAATGSCCATCTCTCTCCAG 6058
QY 6073 gaccgggtaaaattctccttctgtaattcccgccggcgaagtagaanaatgtagaacta 6132
|||||
Db 6059 GACCGGCTAAATTTCTCTTGTCTCAATTCGCGCGGCAAGTACAAATGATGAACCTA 6118
QY 6133 tataagaataagctgctcccaactcagactagattatctaagaataagttcaactga 6192
|||||
Db 6119 TATAAGATTAAGCTGCTCCCATCAGCTTAGATTAATCTAAGAAATTAAGTTTCACTGA 6178
QY 6193 aagaaacaggaataatatacaaaatgtaaaatgtagactggtgggaatagatttcta 6252
|||||
Db 6179 AGGAAACAGAGAAATATTTCAAAAGTGAATAGTATAGACTGGGGAATGTTTTTTA 6238
QY 6253 ttaataagcggggagcagaggtctcacttaaccatcgctccttagatagagaagcgggaca 6312
|||||
Db 6239 TTATATGGCGGGGAGCAAGGCTCCACTTAACTTCGCTTACGATTAAGAGACGGGGACA 6298
QY 6313 gaacccccgtgagcaatgtagaccggaataagtaactgagcgaacaggggccccgcctg 6372
|||||
Db 6299 GAACCCCTGTGGCAATGGGACCCGATTAAGTACGTGCAACAGGGGCCCCCGGCCCTG 6358
QY 6373 gtagcaacgcataactgtgcggtgcccccaataaactcgtcgtcgagctgacaatacaag 6432
|||||
Db 6359 GAGCCACCGCATTAATTTGCGGGTCCCAATTAACTGCTGGGCTGACATTAACACAG 6418
QY 6433 ccgcctagcaacagtaaccactgtgattccttaaccaacgcgcctagaaactcccaagt 6492
|||||
Db 6419 CCGCTTAGCAACAGTACACTGTGATTGATTCACCAACAGCGCTGAGAAACTCCCAAGT 6478
QY 6493 gtcccgtagtaagaagcaagagacacttcagctcatccagggagcttcccaagcactc 6552
|||||
Db 6479 GTTCCGTTAAGACAGGAGACAGACTTCTAGTCTCATCGAGGACCTTTCCAGGCAATC 6538
QY 6553 aactcaacgcacactgactgcaacttctctgtgtgcttgccttctatccctcaagggcctc 6612
|||||
Db 6539 AACTCCACCGACCTGATGCTGCTTCTGTGCTTGTGTATCTCCAGGGCTCTCT 6598
QY 6613 tattataggggagtagcctaagaagaanaatlcactgtgacaagaagacatagaaatcaa 6672
|||||
Db 6599 TATTATAGGGGATGGCTTAAGAAAGAAATTAATCAATGTGACCAAGAGCATGAATTCAA 6658
QY 6673 tgtacatgggggtcccgaaataagcttacccctacactgaagttccggggaaggagacatgc 6732
|||||
Db 6659 TGTACATGGGGGTCCGAAATTAAGCTTACCTCATCTGAAGTTTCCGGGAAGGGGCAATGC 6718
QY 6733 ataggaanaagctcccccatacccaacacacttgcataagtagtctgtgttctgaagcag 6792
|||||
Db 6719 ATAGGAAAGCTCCCATCCCAACCAACACTTGTGCTATAGTACTGTGGTTTATGAGCAG 6778

QY 6793 gccctagaanaatcagtaattagtaacctgtgtatacaagatggtggtgcatgcaatactggg 6852
|||||
Db 6779 GCCTCAAGAAATCAGTATTAGTACTGCTGTTATACAGGTGGGGCATGSCAATACTGGG 6838
QY 6853 ttaacccccgtgttcccaactcagcttcaacaaatccaagaattctgtgtcatgctc 6912
|||||
Db 6839 TTAAACCCCGTGTTCACCTCAGTCTTCAACCAATCCAAAGATTCTGTGTCATGCTC 6898
QY 6913 caaatgctcccgagtgatctacacatcccgaggagtggtctcttgatgtaattgactat 6972
|||||
Db 6899 CAAATCTCTCCCGAGGTACTACATCCAGGAATGGTCTTGATGAAATATGACTAT 6958
QY 6973 cggatataccgaccaaaagaagaccgtaaccttaaccttaactgtaagtcgagatla 7032
|||||
Db 6959 CGGTATACCGACCAAAAGAGAACCGTATCCCTTACCTAGCTGTAATGCTCGGATTA 7018
QY 7033 gggacggccgttggcgttaggaacagggacagctgcccgtgatacaaggaacacagcagta 7092
|||||
Db 7019 GGGAGCGCCGTGGCGTAGAGAACAGGACAGCTGCCCTGATACAGAGACACAGACGCTA 7078
QY 7093 gagaagaagactgtgtgagctacatgagcgcaatgacagaagaatctcagagccttaagag 7152
|||||
Db 7079 GAGAAAGGACTTGGTGAAGCTACATGCGGCCATGACAGAAAGATCTCCAGCTTAAGAG 7138
QY 7153 tctgttagcaacctagagaagatccctgacttctgtcgtgaagtgtctacagaacgg 7212
|||||
Db 7139 TCTGTAGCAACCTAGAGAGTCCCTGACTTCTTGTGCTGAGAGGTTCTACAGAACCGG 7198
QY 7213 agggatatagtctgctgttctlaagaagaagtgtgtatgtgtagccttaaaagaaga 7272
|||||
Db 7199 AGGGATTAAGATCTGCTGTTCTAAAGAGAGTGGTTATGTGACAGCTTAAAGAGAA 7258
QY 7273 tgtgtctctatgtatagtaactcagagggcaatgaagacccctgaacaaagcttagaaa 7332
|||||
Db 7259 TGTGCTTCTATGTAGATCACTCAGGAGCCATGCAAGACTCCAGAACAGCTTAGAA 7318
QY 7333 aagttagaagagctcaggaaggaagaagcctacacaggggtgtgttgaagaatgtctc 7392
|||||
Db 7319 AAGTTAGAGAGCGCTCGAAGGAAAGAGGCTGACCAAGGGGTGTTGAAGATGGTTC 7378
QY 7393 aacagctcctctgtagtaccacccctgctctcgtcctgacggggccccctagtagtctg 7452
|||||
Db 7379 AAGAGTCTCTTGGATGACCACTGCTGCTGCTGAGCGGGGCCCTAGTGTCTCTG 7438
QY 7453 ctctgttacttaacagttgggcttgccttaataatagttgtgtgcttgttaagaa 7512
|||||
Db 7439 CTCTGTACTTACTAGTTGGCGCTTGTCTTAATATAGTTGTGCTTGTATGAGAA 7498
QY 7513 cgaagtgtgcaagtcagatcatatgtaacttaggcaacagtaaccaagcctctgaagcaa 7572
|||||
Db 7499 CGAGTGAAGTGAAGTCAATATGATGTTAGGCAACAGTACCAAGGCTCTGAGCCAA 7558
QY 7573 ggaagaactgacactcctcctccagcttcaagatagaactlaaacaagaagaag 7632
|||||
Db 7559 GGAGAAACGAACTTACCTTCCAGTTCAAGTATGAACTATTACAAGACAAGAG 7618
QY 7633 tggggaatgaaagatgaataatgcaactaacctccccaagaacccaagaagttaataaa 7692
|||||
Db 7619 TGGGGAATGAAGAGATGAAGATGAACCTTACCTCCAGAACCCAGGAATTTATTAATA 7678
QY 7693 agctctaagtgcccccgaatcaagaacctcgtgctcgaatgaatagtagaagtgca 7752
|||||
Db 7679 AACTCTAAATGCCCCGAAATTCAGAACCTGCTGCTGCTGCAAGTAATAGTAAAGTGCA 7738
QY 7753 caactcctatgttcccaagggcctgctatccctgagccttaagtaagaatacagaanaatgagtt 7812
|||||
Db 7739 CACTTCCATATTTCTCAGGGCTGCTATCTGAGCTTAGTAAGATTAACAGAAATGAGTT 7798
QY 7813 gactaaatcgtcttaactgtatctgttaaaactgactgagcaacatagaaagaattgattac 7872
|||||
Db 7799 GACTAATGCTTATCTGATGATCTGTAAACAGTACAGGACCATTAAGAAATGATTAAC 7858
QY 7873 attgacagccctagtagactatacactgcaactgtcactcgtccagagagccacgc 7932

Db	7859	ATTACACGCCCTATGTGACCTATCTCCAACTGCATCTGTCTACTCTGCCAGAGCCACGC	791.8
QY	7933	AGATGCGAGCTCCGGAGCGATTTTAAATGATGTGTCACAGCAGCGGGCTTCGATA	799.22
Db	7919	AGATGCGAGCTCCGGAGCGATTTTAAATGATGTGTCACAGCAGCGGGCTTCGATA	797.8
QY	7993	TTCTAAATGATGTGTCACAGCGCGGCTCGATATTTTAAATGATGTGTG	805.2
Db	7979	TTTAAATGATGTGTGTCACAGAGCCGGGCTTCGATATTTTAAATGATGTGTG	803.8
QY	8053	ACGCACAGGCTTGTGTGTAAGCCCAATAAAGCTGCCGATCCGCACCTCGGGCGCA	811.2
Db	8039	ACGCACAGGCTTGTGTGTAAGCCCAATAAAGCTGCCGATCCGCACCTCGGGCGCA	809.8
QY	8113	GTCCTCAACCCCGCGTGTGTGTCAGACTGTGGGCCCGAGCGGCTTGGAATAAAATCT	817.2
Db	8099	GTCCTCAACCCCGCGCTGTGTGTAAGCACTGTGGGCCCGAGCGGCTTGGAATAAAATCT	815.8
QY	8173	CTTGCTGTTCATCAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG	820.9
Db	8159	CTTGCTGTTCATCAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG	819.5

RESULT	3
LOCUS	PEN293657
DEFINITION	PEN293657 8763 bp DNA
ACCESSION	Porcine endogenous type C retrovirus proviral gag gene, pol gene
VERSION	and env gene, class B, clone 213.
KEYWORDS	AJ293657 AJ293657.1 GI:14275841 env gene; envelope; gag gene; group specific antigen; pol gene; polymerase; polyprotein. porcine endogenous type C retrovirus.
SOURCE	

FEATURES	Location/Qualifiers
source	1. .8763

```
/translation="MGQVTTPLSLTDHNTVYKSRAHNLVQYKKGPWQTFCASEWPTFDVGWPESEGFNSEILLAVKAILFQTPGSHDQEPYLLTWQDLAEPPPWVKPWLNLKPRKPGPRILALGEKNKHSAEKVEPSPGIYETIEEPTWPEPQVPPPPYPQAQAVRG
```

gene
 CDS
 <2652..6239
 /gene="pol"
 <2652..6239

```

/ gene="pol"
/ note="complete pol ORF, gag-pol is expressed as a
polyprotein"
/ citation={1}
/ codon_start=1
/ product="polymerase"
/ protein_id="CAC39626.1"
/ db_xref="GI:14275843"

```

/leasatla1cion="GRCSIDPLPERATYALKEGSPVEFLVDZAGHNSVLLLOGLPKGLE
 KKSWMATGICOROMPTTRVTVLDVAGRVHLSVIEPCQVPLILGRVTLKMCANOSI
 EGOREVSVNCSMTPTVLTLOLDEYRKLXSPQVCPDIOISMLDFOPOANAETKMGCLA
 KOVPOVYOLKASAFPVSVRYOYPSRABRBCIMPHVOURLQOGLVLYVQSPMNPPLP
 VRKGTGNDYRPVODLREYKRVODIHPTVAPVNLISALPERKWTVLIDUKAFNPLP
 RLHPHSOPLFEHEDPGTGTEGQTLMTPLBQGCNSPTEFEDLHMDLANFTIOHPO
 VLTLOQYVNDLLAGATKODCELEGTAKALLREDSIGYASAKKAOICREATVYLSIGR
 GGOSELEHARKKVFVOPAPPTAKOYREVEFGTGEQSLMPTGATVAPLPYPLTKEGE
 EFSNAPEHOKAFVDAIKKALLSAPALAPADVTKEPTIVEDSKCARSVLEPOTGPMR
 PAVATSKLIDPVSAGMCPVCLKAIAVAALIVKADKIDILGONITVIVPAHLENTYVOP
 DPMVTKNMTJHVOSSLITERVTEPAPALAPADITLPERPDEPHOCHOLITERVYR
 KLDLIDPGEVLVMTPTDSSVYVEEGKMGANVOSTRTIANSIPEBTSOAKETLM
 ALUOLDILRABESKINITYTDSRYAPATANHGAITYKOGULITSRGREKINKEETLSLE
 ALHPKRLAIIHCGHOKAKDLISRGOMQMDRVAKAOQAVNLLPIETEPKAPPEPPO
 YTLDEWOCIKIDGOSTSETPECTCYTSLGKELIPAKLEVEVOQIHRITLNGVTHNLO
 VTRISVYHVLRLPGVADSVVHCVCPOLVANNPSRIPBEGLKSRSHSGAHNEVDFTVFK
 PAKGNGVYLLVFEVDTPESGWEAYPTKKESTVAAKILIEIPREPIKVIYSGDNCPA
 FVAGOSGLKILGIDIMKRLICAVROSGSGOVEMNRIKETIKLITTEINGIMVAIL
 PVEYLERANTPGOGLPYTELVLGGPPELVEIIVASVADALISOPLESFRLKALEWRO
 RAMKOLKEVSGEEDLOVPHRFOVGVSVYRRHHAANDLERMKGSPVILVITPTAVKY
 EGISTMIIHSHSVKLAPPDSDGMAKERTENLKLRLHRLVYSNNNSPO"
 6113 8085

CDS

LTR
 polyA_signal
 COUNT 2301 a 2200 c 2287 g 1974 t 1 others

```

Dry Match          97.9%;  Score 8039.8;  DB 14;  Length 8763;
t Local Similarity 99.2%;  Pred. No. 0;
Chs 8132;  Conservative 0;  Mismatches 58;  Indels 8;  Gaps 5;

```

1 gtagtgtagactgtgagccacgacgccttgataataaacctctctgtgttcac 60
|||||
490 GTGGTATACGACTGTGGGCCCCCAACGCGCTTGGAAATAAAATCTCTGTGTTTCATC 549

QY 61 aagacgcttcctgtagtgatattgggtgtcgctcttcgagcccgagagggagat 120
|||||
Db 550 AAGACCCCTCTTGTGATGATTTGGGGTTCGCTCTCCGAGCCGAGAGGGGAT 609
QY 121 tttctttactgaccttaatttgggtgtggttgccgggaaatccctgcacccctta 180
|||||
Db 610 TGTCTTTTACGTGGCCCTTCAATTTGGGTGGCTGGGAAATCTCCGACCAACCCCTTA 669
QY 181 caaccggaacccgaccttgaggttaagagatccaccttgaaacgtgtgtgtcgagcc 240
|||||
Db 670 CACCCGGAACCCGACTTGGAGGTAAGGATACCCCTTTGGAACGTGTGTGTGTGGCC 729
QY 241 ggcgtctgttctgtagtgtctgttctcgatgtgcgacttcogtlttgcaactgtcct 300
|||||
Db 730 GGCCTCTATGTTCTGATGATGCTGTGTTCCGTAATGCGCGCTTCCGTTTSCACTGTCCT 789
QY 301 ctacagacctaagagactgtagagactgtatcaacagacgtgtctagagagatacaagctg 360
|||||
Db 790 CTCAGACCGTAAGACTGGAGAGACTGTGATCACAGACGTTTAAAGAGATCAACAGGCTG 849
QY 361 ccaacctggggagaccccgaggaggttgaggagagccagagacgctgtgtctctact 420
|||||
Db 850 CCACCCGAGGGAACGGCCCGGAAGGTGGGAAGCCAGAGAACCCGTGTGTGTCTCTACT 909
QY 421 gtcgttagagagagacgagttctgttltgaagcgaaagcttccacctcogcgcgctcg 480
|||||
Db 910 GTCGCTGAGAGGTCCGAGTTCTGTGTTGAAGCGAAAGCTTCCCTCCGCGCGCTCGCG 969
QY 481 accttttgcctctgttgaaagagcgagcggtgcggtgtgtctggatctgttgttt 540
|||||
Db 970 ACTCTTTTGCCTGTGTGGAGAGCGGGAGCGGGCTGCTGTGTGATCTGTGTGTGT 1029
QY 541 ctgtctcgtgtgtcttctgtctgtgtgtgtctgttctcaagtttaatatagagacag 600
|||||
Db 1030 CTGTTGT 1089
QY 601 tgaatacccccttagtttgaactctcgacattgagctgaagtlagatccagggctcata 660
|||||
Db 1090 TGACTACCCCTTGTGATTTGACTCTGACCATTTGAGCTGAACTTGAATCCAGGCGCTCA 1149
QY 661 atttgacgttcaggttaagaaaggaccttgacagacttctgtctctgaaatgagcga 720
|||||
Db 1150 ATTGTGTAGTTCAAGTTTAAGAAAGGACCTTGGCAGACTTGTGTGTGTGTGTGTGTGT 1209
QY 721 catctgactgttgatgagccatcagagagggagaccttaattcgtgaatctcgtgtgta 780
|||||
Db 1210 CATTCGATGTTGATGGCCATCAGAGGGGACCTTTAATTTCTGAATTTATCTTGCTGT 1269
QY 781 aggaacatcatlttccagactgtagccggtctcatccctgatacagagccctatalactta 840
|||||
Db 1270 AGGCAATCATTTTTCAGACTGGACCGGGCTCTCATCTGTATCGAGAGCCCTTATCTT 1329
QY 841 cgtggcaagatttgagagaagatccctcgcgcacatgggttaacacatgtgtctaaataacaa 900
|||||
Db 1330 CGGTGGCAAGATTGGGCAAGAAACCTCCGCAATGGGTTAACCATGTGCTAAATAACCA 1389
QY 901 gaaagcagagttcccgcaatcctgtgtctgtgagagaaaaaacaacactgcggccgaaag 960
|||||
Db 1390 GAAAGCGAGGTCCCGCAATCTGTGCTTGGAGAGAAAAACAACACTGGCGGAAAAAG 1449
QY 961 tgaagccctctctcgtatctaccccgagatcgaagagccgcgacttgcgcgagacc 1020
|||||
Db 1450 TCGAGCCCTTC-TCTGTGTATCTTACCCCGAGATGAGAGACCGCGCACTTGGCGGAA 1508
QY 1021 caacctgttcccccaccccttaaccagacaggtgtgtgtgtgtgtgtgtgtgtgtgtgt 1079
|||||
Db 1509 CAACCTGTTCCTCCACCCCTTATTCAGACAGAGGTGTGTGTGTGTGTGTGTGTGTGTGT 1568
QY 1080 ccttgagactcaggttgtagaggaactgtgtgcgagacccgagccgagagagcgccacc 1139
|||||
Db 1569 CCTGAGACTCCGGTGTGGAGGAGACTGTCTGCCGGGACTCGAGCCGGAGAGAGCGCGCAC 1628
QY 1140 ccgagagcagacagagatcgcatattacgcgtgtgcacactatgtgcccctccatgtcca 1199

Db 1629 CCGSAGCGGACAGACAGATTCGATATATACCGCTGCGGCACTATGAGCCCTCCATGCA 1688
QY 1200 gggggccaattgcaagccctccagatattgagccctttcttcgaagatctcataattgg 1259
|||||
Db 1689 GGGGGCAATTGCAAGCCCTCCAGTATGTGGCCCTTTCTTCTGTGAGATCTCATTAATGG 1748
QY 1260 aaactaaacatcccccctctcagagagatcccaacgctcaaggggttggtagagctc 1319
Db 1749 AAACATACACATCCCTCTTCTTCGAGATATCCCAACGCTCACGGGGTGTGTGAGTCC 1808
QY 1320 ctatgttctcacaagacctactcttgagatgtgttcaacagctgtctgaacacactctc 1379
|||||
Db 1809 CTTATGTTCTCTACACAGGCTACTTGGGATGATTTGCAACACGCTGCTGACACTCTTC 1868
QY 1380 acaacgagagacgagagagaaattctgttaagagcttagaataaaatgttccctggggcgac 1439
|||||
Db 1869 ACAACCGAGAGACGAGAGAAATCTGTAGAGGCTAGAAAAAATGTTCTTGGGGCGGAC 1928
QY 1440 gggcgacccacgagltgcaaatagatgtgacatggagatlttcccttgactgcgcccgt 1499
|||||
Db 1929 GGGCGACCCACCGAGTTGCAAAATGAGATGTGATGGGATTTCTTGTACTGCCCCGCT 1988
QY 1500 tgggactacaacacggtctgaaggttagagagcttgaatctatccgcaagctctgtgt 1559
|||||
Db 1989 TGGGACTACAAACAGGCTGTAAGGTAGGGAGAGCTTGAATAATCTATCGCAGGCTGTG 2048
QY 1560 ggcggtgtctcggggcgccctcaagagcgccactaaatttggctaaagttagagaggtgt 1619
|||||
Db 2049 GCGGGTCTCGGGGGCGCCCAAGACGGCCCACTAATTTGGCTTAAAGAGAGGTGATG 2108
QY 1620 cagagacccgaacgaactccctcgttattctctgagagctcatatgaagcccttcagcg 1679
Db 2109 CAGGGACCGAAGCAACCTCCTCGGTATTTCTTGAAGAGCTCATGGAAGCCCTTCAGCGG 2168
QY 1680 ttcaaccccttttgatcttaacctcagaagcccaagaaagcctcaagltgcccctgtctca 1739
|||||
Db 2169 TTTACCCCTTTTGTGATCTTACCTCAAGAGGCCCGGAAGGCTCAGTGGCCCTTTCAT 2228
QY 1740 gggcaggtgcgtctgatatatacagaaactctgaagactggaaggttcaagaagct 1799
|||||
Db 2229 GGGCAGTGGCTCTGTGATATCAGAAAGAACTTCAAGACTGGAAGGGTTCAGAGAGCT 2288
QY 1800 gagttacgttgatctagtagagagagcagagaaggtgtatccaagagagaaacaaag 1859
Db 2289 GAGTTACGTGATCTAGTGTAGAGAGGCAAGAAAGGTGTATTACAGAAAGGAGACAGAG 2348
QY 1860 gagaaggaacagagaaaaaaggaagagagaagaaaggagaaacgtgtatagacgt 1919
Db 2349 GAGAGGAACAAGAGAAAGAAAGAGAGAGAGAAAGGAGAAAGAGCTGATAGACGG 2408
QY 1920 caagagaagaatttgactcaagatcttgcgcagatgtgtgtgaagaggaagcaagcagggag 1979
|||||
Db 2409 CAAGAGAAAGAAATTTGATCAAGATCTTGTGCGCAGAGGTTGAAGGGAAGAGAGAGGAG 2468
QY 1980 aagagagagattttagaataaattaggtcagagccctaaacgctcaaggaagactggcaat 2039
Db 2469 AGAGAGAGAGATTTAGGAAAAATTAGTCAAGCCCTTAACACGTCAAGGAACTTGGGCAT 2528
QY 2040 aggaacccactcgaacaaagacagltgtcgtatgttaagaagaaaaagacacttggcaag 2099
|||||
Db 2529 AGGACCCCACTCGACAAGACAGAGTGTGCTATTGTAAAGAAAAAGGACACTGGGCAAG 2588
QY 2100 aactgtccccaagagaggaacaaagagacccaaggtctcagtctcagaagaagataaagat 2159
Db 2589 AACTGCCCAAGAAAGGAAACAAAGAGACCGAAGGCTTCAAGCTTCAAGAAAGATTAAGAT 2648
QY 2160 tagggagacgaggttgcagacccctcccgagcccaaggggttaacttgaaggttagaggg 2219
|||||
Db 2649 TAGGGGAACCGGGTTCGGACCCCTCCCAAGCCCAAGGTAACTTTGAAGGTGAAGGGG 2708
QY 2220 caacagttgagttcctgtgtatgatacaggaagcagacatcagtgctgtctacaaccta 2279
|||||

Dh 2709 CAACAGTGTGATCTGTTGATACCGAGCGAGCATTCAGTCTGTACAAACATTA 2768
Qy 2280 ggaactaaagaaagaaatctctgggtatgggtgccaagggcaagcgatcca 2339
Dh 2769 GGAAGAACTAAAGAAAAAATCTGGGTGATGGGTGCGACAGGCAACGGCATATCA 2828
Qy 2340 tggactcccgaaagaaacgttgaacttggagtggaagtgaaacccctgttctgtc 2399
Dh 2829 TGGACTACCCGAAAGACCGTGGACTGGAGTGGAGCGGGAACCCACCTGTTCTGTGC 2888
Qy 2400 atcccttagtgccagtaacccctcttagtgtagagacttaactgaacaaatggagctca 2459
Dh 2889 ATCCCTGATGTCCAGTACCCCTTGTAGTGAAGACTTGTGACCAAGATGGAGTCA 2948
Qy 2460 attcctttgacaaggaagccagagtgctgtgtaatacaaacccactgctgtg 2519
Dh 2949 ATTTCTTTTGAACAAGAGACCGAAGTGTCTGTGAATACAAACCCATCCTGTGTG 3008
Qy 2520 accctcaatagaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 2579
Dh 3009 ACCCTCAATTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3068
Qy 2580 atacaagctgt 2639
Dh 3069 ATACAGTCTGT 3128
Qy 2640 gcaagaagaagtlcccccacaggtatcaactgaagcagtgctacacagatcaatc 2699
Dh 3129 GCAAGAGAGTGTCCCGACAGTATTCATGAGGCGCATGTATACAGATATCATGTC 3188
Qy 2700 agacagtaaccccttgatgaagagagctcgagaagaatttggccgcatgttcaaatga 2759
Dh 3189 AGACAGTACCCCTTGTAGTGAAGAGGCTCGAAGAGATTTGGCGCATGTGTCAAGATTA 3248
Qy 2760 atccaagaagagcatccctgaatccctgttccaaatcccttgatgaatccctgttccaa 2819
Dh 3249 ATCCAAAGAGGATCTTGT 3308
Qy 2820 aggaagccttggaacaaatgaatgaacagtaacaggaacttgaagagtgtaataaag 2879
Dh 3309 AGGAAGCCTGGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3368
Qy 2880 gtgtagagacatacaacccaaaggttccgaacccctataactcttggagcgccctccgct 2939
Dh 3369 GTGAGAGACATACACCCAGAGGTCCGACACCTTATTAACCTTGAACCGCCCTCCGCT 3428
Qy 2940 gaaaggaagtgatcacagatgaacttaagaatgctctctctcgcgagatlaaac 2999
Dh 3429 GAAGGAAAGTGTACAGATTTGGACTTAAAGATGCTTCTTCTGCTGAGATTTACAC 3488
Qy 3000 ccaactagccaacacttttgccttcgaatgaagagatccaggtacgggaagaacgg 3059
Dh 3489 CCCACTGAGCCAACTTTTGTGCTTGAATGAGATGAGATGAGATGAGATGAGATGAGATGAG 3548
Qy 3060 cagctaaccttggaacgagctggcccaaggttgaagaactcccgccactcttggagaa 3119
Dh 3549 CAGCTCACCTGTGAGCCCACTGCCCCAAGGTTCAAGAACTCCCGAATTCCTTTGAGAA 3608
Qy 3120 gccctacaagagaccttgcgaacttcaagatcaaacacccctcaggtgacccctccag 3179
Dh 3609 GCCCTACACAGGAGACCTGGCACTTCAGATCCAAACCCCTCAGGTGAGCCCTCTCTCAG 3668
Qy 3180 taagtgatgaactgtcttgcgaggaagccaacaaaggaactgtctagaagtlagaag 3239
Dh 3669 TACGTGATGACCTGCTTGTGGGGGAGCACCAAGAGAGTGTAGAGGTAGAGAG 3288
Qy 3240 gcaactatgctgtgatatgtctgacactaggtacagagccctgtctaaagagccagatc 3299
Dh 3729 GCACTACTGTGGAATGTGTGACCTAGCTACAGACCTCTCTAAGAGGCCAGATT 3788
Qy 3300 tgcagagagagtaacatacttgggttacagtttgcgggagggagagagatgtgcagc 3359
Dh 3789 TGCAGGAGAGAGTAACTACTTGGGGTACAGTTTGGGGGGGGGAGCATGGCTGACG 3848
Qy 3360 gaagcaggaagaaactgtatccagataccggccccaacacagccaagtgaga 3419
Dh 3849 GAGGCGAGAGAGAAACTTACTCCAGATACCGGCCCCACACAGCCAAACAAATGTAGA 3908
Qy 3420 gaatttgggagacgttgatcttgcagactgtgatcccggttggcagacttaaga 3479
Dh 3909 GATTTTGGGACACTGATTTTGCAGACTGTGATGCCGGGGTTGGCAGCTTAGCA 3968
Qy 3480 gcccacttaccggcttaacaaagaagggatctctgtgctcctgagacag 3539
Dh 3969 GCCCACTTACCCGCTTAACCAAGAAAGGGAATTCCTGGGCTCTGAGACAG 4028
Qy 3540 aaggaatttgatgatacaaaaagccctgcgagcagcagcagcagcagcagcagcagcagc 3599
Dh 4029 AAGGCTTTGATGATATCAAAAAGCCCTGTCAGACGACCTGCTGCTGCTGCTGCTGCTGCT 4088
Qy 3600 gtaactaaaccccttacccttgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 3659
Dh 4089 GTACTAAACCTTTTATCCCTTATGTGATGAGCTTAAGGAGTGAAGCCGAGAGATT 4148
Qy 3660 acccaaaccttagagacatgagagagacgtgtgctactgttcaagaagcttact 3719
Dh 4149 ACCCAACCTTAGAGACATGAGAGACCTGTGCTTACCTGTCAAGAGCTGATCT 4208
Qy 3720 gtagccagtggttggccgatagtlctgaagagatcagcagtgtagcactgtgtcaag 3779
Dh 4209 GTAGCCAGTGTGTGGCCGTATGCTGAAGCTTATCCAGCTGTGCTGCTGCTGCTGCTGCT 4268
Qy 3780 gacgttgacaatgaacttggagacaaatgaactgttaatgaatgaatgaatgaatgaatgaat 3839
Dh 4269 GACGCTGACAAATTTGATTTGGAGACGAATATATATATATATATATATATATATATATAT 4328
Qy 3840 aacatgttgcgacagcccccagacagatgaatgaatgaatgaatgaatgaatgaatgaatgaat 3899
Dh 4329 AACATGTTTGGGACGCCCAAGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 4388
Qy 3900 agcgtgtcttcaacagagaggttcaacttgcctcacaacagcgcgtcacaacccctgac 3959
Dh 4389 AGCCTCTTCTCAAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4448
Qy 3960 ctctgtcctgaagagatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 4019
Dh 4449 CTCTGCTTGAAGAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4508
Qy 4020 ggaacttggttgcgaagacacttcaacacacacacacacacacacacacacacacacacacac 4079
Dh 4509 GAGACTGGGGTCCGAGAGACCTTACAGACATACCGCTGAGCTGAGAGTGTGTAACCTGG 4568
Qy 4080 ttaacttgacgaagacagatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 4139
Dh 4569 TTCACAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4628
Qy 4140 gacggagacccgagacatctgggacagcgcgtgcggaaggaacttgaagcgcaaaagct 4199
Dh 4629 GAGGGAGCCGACAGATCTGGGGCAGACGCTGCGGAGAGAACTTCAAGCCGAAAGGCT 4688
Qy 4200 gagctatgacctcagcaagcttgcgctgcggaaggaatcaataaacttact 4259
Dh 4689 GAGCTATGCGCTCAGCGAAGCTTGGGGCTGGCGGAGGAATTCATTAACCTTAT 4748
Qy 4260 acggaacagatgatgcttgcgactgcagcgtacacggtacacggtggtggtggtggtggtggt 4319
Dh 4749 ACGGACAGAGGTATGCTTTCGACTGACACGCTACACGCGGCGCATCTATTAACAAAGG 4808
Qy 4320 ggggttgctactcagcagggaggaataaagaacaaagggaattcttaagctata 4379
Dh 4809 GGGTGTCTTACCTCAGCAGGAGGGAATTAAGAAAGGAAGGAATTTCTAAGCTTATTA 4868
Qy 4380 gaagccttacttgcgaagagcttacttatacacttccctgagacacgaagacc 4439
Dh 4869 GAAAGCTTATTTGCAAAAAGGCTTACTATATATATATATATATATATATATATATATATATAT 4928

Query Match	Best Local Similarity	Score	DB	Length
Matches 8124: Conservative	99.14;	Pred. No. 0;	Mismatches 54;	Indels 20; Gaps 6;

QY	1	gtgtgttaaacactctgtgagcccaagccgcctctgtaataaaataacccctctgtcttgtaac	60
Db	489	GTGTGTATACACTCTGTGGGCCCCACGGCGCTTGGAAATAAAATTCCTTCTGTTTGATC	548
QY	61	aagaccctctctgtgtgatttctgtgtgtctgtccctctcagaccgcagaggggatt	120
Db	549	AAGACCGCTTCTGTGATGATTTTGGGGTGTGCCCTTTCGAGCGCCGGACGAGGGGAT	608
QY	121	tgctcttctaactgtgcttcatctgtgtgtgtgtgtccggagaaactctgacacccctta	180
Db	609	TGTTCTTTTACTGTGCGCTTTTCATTTTGGGTGCTTGTGGCGGAAACCCGTCGACCAACCCCTTA	668
QY	181	caaccggagaaacagcactgtgggttaagggatccccctttgaaacgtgtgtgtgtgcgac	240
Db	669	CACCCGAGAAACCGACTTGGAGGTAAAGAGATCCCTTTTGGACGTGTGTGTGTGTGGGCC	728
QY	241	ggcgtccctgtctctgaagtgtcgtcttttcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	300
Db	729	GGCGTCTCTGTTCTGAATGTC-----TTGGCGCTTTCGGTTTGGACGTGTGT	776
QY	301	ctcagaccgttaaggaactctgtagaactgtatcaagcaagctgtcctaagagatcacagactg	360
Db	777	CTCAGACCGTAAAGACTGTGAGGACTGTGATCAGACACGCTGTAAGAGATCATACAGCGTC	836
QY	361	ccaccctgtgggtgacgccccgggaggtgtgtggagagccagggagcgtgtgtctctact	420
Db	837	CCACCCCTGGGGGAGCGCCCGGGAGGTGGGAGAGCCAGGAGACCGCTGGTGTCTCTACT	896
QY	421	gtcgctcagaaggaacaggaattctgtgtgtgaagagaaagctccccctccgcgcgcgtcg	480
Db	897	GTCGCTCAGAGACCGAGTTCTGTTGTGAAGCGAAAGCTTCCCTCCGGGGCGGTCCG	956
QY	481	acctctttgacctgtctgtgaagaacgcgcgagctgcgcgtgtgtgtgtgtgtgtgt	540
Db	957	ACTCTTTTGCCTGCTTGTGGAAGACGGGAGCGGGTGTGTGTGTGTGTGTGTGTGTGT	1016
QY	541	ctgtctcgtgtgtctctgtctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	600
Db	1017	CTGTTTGGT	1076
QY	601	tgactaccccccttaagttaacctctgacatctggaactgaatlaaactcagaagctcata	660
Db	1077	TGACGACCCCCCTTAGTTTAACTCTGACCATTTGGACTGAAGTTAATCTCAAGGCTCAT	1136
QY	661	attctcagttcaggttlaaagaaggaacctgtgcagacttctgtgtcctctgtatgtgcaaa	720
Db	1137	ATTGTTCAGTTTCAGTTTAAAGAAGGAGACTTGGAGACTTTCGTGTGCTCTGAATGAGCCA	1196
QY	721	catcogaatgtgtgattgtgcactcagaagggagccttlaactctgaataatctgtcgtgta	780
Db	1197	CATTGCATGTTGTGATGTGGCCATACAGAGGGGACCTTTAATCTGAAATTAATCTCGCGCTGTA	1256
QY	781	aggaataatattttcaaacactggaacccgcctatactctatcagaagagccctatactcata	840
Db	1257	AGGCATATATTTTTCACACTGTGAGCAACGCTCTATCTGTATCGAGGAGCCCTAATATCTTA	1316
QY	841	cgtgcgaagattttgcagaagaatctctccgcacatgtgttlaaacatctggtctaaataaacaa	900
Db	1317	CGTGGCAAGACTTGGCAGAAAGATCCTCCGCATGGGTTAAACCATGGTTAAATTAACAA	1376
QY	901	gaaagccaagttccccaatctcgtgctctctgtgagagaaaaacaacaactcgcgcgaaaaag	960
Db	1377	GAAAGCCAGGTGCCCGAATCTGTGCTTGTGAAGAAAAAACAACCTCGGGCGAAAAAG	1436
QY	961	tcgagccctctctgtatctaccccgagatcgaagaacgcgcgcgcgcgcgcgcgcgcgcgc	1020
Db	1437	TGAGACCTTC-TCTCTGATATTCGCCGGAATGAGAGAGCGCGCGACTTGGCGGAGACC	1496
QY	1021	caaacgtgtctccccaacccttataccagacagagtgctctgtgtgagga--cctctgcacct	1076
Db	1496	CAACCTGTGTCCCAACCCCTTATCCAGCAGACAGAGGTGCTGTAGGGGAGACCTCTGCCCC	1556
QY	1080	ctctgagactccggtgtgtgagaggaactgtctccgggaactcogaaccggaagagcgcaac	1136

QY 3300 tgcagagagagatgaataacttgggttacagtttgcgggcgggcagatggctgaag 3359
| | | | |
Db 3776 TGcAGGAGAGAGGTAAcATACTTGGGTACAGTTGGGGGGCGGCGACCATGGCTGACG 3835
| | | | |
QY 3360 gaagcagagaaataactgtatgcagataccggcccaacacagccaagaatgtaga 3419
| | | | |
Db 3836 GAGGcAGGAGAAAGAAcTGTAGTCCAGATACCGGCCCAcCACAGCCAAAcAAGTGA 3895
| | | | |
QY 3420 gaattttggagacgtgatatttgagactgtgatacccggtgttgagacttagca 3479
| | | | |
Db 3896 GAGTTTtGGGAGAGcTGGATTTTGGAGcTGGATCCGGGgTTTGGACcTTAGCA 3955
| | | | |
QY 3480 gccccacttaaccgcctaaccaaaagagagatcttcctcggtccttgagaccag 3539
| | | | |
Db 3956 GCCcACTCTAACCCGTAAcCAAAAGAAAGGGGAATTCTcTGCGGTCTCTGAGcACAG 4015
| | | | |
QY 3540 aagcatttgatctatcaaaaagccctgtcgaagcactgtccttgccctccctgac 3599
| | | | |
Db 4016 AAGcATTtGATGCTATCAAAAGGcCTCTGAGcGCACCTGCTGCGcCTCCcTGAC 4075
| | | | |
QY 3600 gtaactaaaccccttaacccttatttgatgagcgttaagagtagcccgaggaattta 3659
| | | | |
Db 4076 GTAAcTAAAcCTTTAcCCTTTATGTGATGAGcGTAAgGAGTAgCCGAGGAGTTTAA 4135
| | | | |
QY 3660 acccaaacccctaggaacatggaagagacactgttgcctacactgtcaagaagcttgact 3719
| | | | |
Db 4136 ACCcAAAcCCTAGGACCATGAGAGAGAcCTGTtGCCACTGTcCAAAAGAcTTGATCC 4195
| | | | |
QY 3720 gttagcaggtgttggtcccgatgtctgaagcgtatcgacgtgtgacatactgtgcaag 3779
| | | | |
Db 4196 GTtAGCCAGtGTGGTGGCCGTATGCTGAAGcGTATGcCAAGcTGTGcCATAcTGcTCAAG 4255
| | | | |
QY 3780 gaecgtacaacattgacttttggagacagaataactgaataagcccccaatgatatgag 3839
| | | | |
Db 4256 GAGcGTcAACAATTGATTTGGGAGAcAATATAcTGTATAGcCCCCATGcATTGGAG 4315
| | | | |
QY 3840 aacatcgttcgcgaagcccccagacagatgataccaagcccgcatagacccactatca 3899
| | | | |
Db 4316 AAcATcGTTcGGcAGCCcCAGAcCCATGATGATGACCAAGcCCCGcATGAcCCATTCAA 4375
| | | | |
QY 3900 agcctgcttctcaagagagaggttcaactttgcctcaaccagccgcttcaacctgcact 3959
| | | | |
Db 4376 AGcCTGcTCTCAcAGAGAGGcTCAcTCTGCTCAcAGCCcCTCTCAAcCCTGcCACT 4435
| | | | |
QY 3960 ctctgcctgaagagactgtgaacagtgactcatgattggcatcaactatgattgag 4019
| | | | |
Db 4436 CTtCTGcTGAAGAGAcTGAACCAAGTAcTcATGATTGcCATTAcTATTGATTGAG 4495
| | | | |
QY 4020 ggaacttgggttccgcaagaccccttacagacataccgctgaacttggagaatgtctaacttg 4079
| | | | |
Db 4496 GAAcACTGGGgTCCGCAAGGAcCTTAcAGAcATAcCGcTGAcTGGAAGTGTcTAACCTG 4555
| | | | |
QY 4080 tttaactgaacgaagcagctatgtgtggaagtgaagaatgctgtggtgcgagctgtgtg 4139
| | | | |
Db 4556 TTAAcTGAcGGAAGAcAGcTATGTGTGAAGGTAAAGAGATGGcTGGGcGGGTGGG 4615
| | | | |
QY 4140 gacggagaccgcacgaatcttgggcagcagcctgcggaaaggaactcagcgcaaaagct 4199
| | | | |
Db 4616 GAGGGAAGCCAcACAGATGTGGGcCAGAcCTGcCGGAAGGAATTCTAcCGCAAAAGcCT 4675
| | | | |
QY 4200 ggaactatggcctcaagcaagcttgggcgttggccgaaaggaatcatalaacttlat 4259
| | | | |
Db 4676 GAcCTATGGCCCTCAcGGAAGcTGTGGcGTGGcGGAAGGAATTCATTAACATTTAT 4735
| | | | |
QY 4260 aaggaagcaggtatgacttttggagactgacacgttacaagggccatcataaagaag 4319
| | | | |
Db 4736 AcGGAAGAcAGATATcCTTTGGAcTGCACAGTAcAGGGGcCATCTATTAACAAAG 4795
| | | | |
QY 4320 ggggttctaactcagcagaggaagaaataaagaacaagaaggaattctaagctata 4379
| | | | |
Db 4796 GGGTTcCTTAcCTCAcCAGAGGAGGAATTAAGAAcAAAGAGGAATTCTTAAGcCTATTAA 4855
| | | | |
QY 4380 gaagccttaacatttgcacaaaaggtagctatattatacactgtcctggacatagaagcc 4439
| | | | |
Db 4856 GAAcCTTAcATTtTGGCAAAAGGcTAGcTATTATATAcTGTCTGcCATATAGAAAGCC 4915
| | | | |
QY 4440 aaagatcataatcagaaggaaccagatgtcgtgacgggtgtgccaagagcagccag 4499
| | | | |
Db 4916 AAAGATCTcATATCTAGAGGGAACcAGATGGcTGAACCGGGGTtGCCAAGAcGACCCAG 4975
| | | | |
QY 4500 gctgttaacctctgccttaataagaaccccaagcccaagacccagacagcagtagc 4559
| | | | |
Db 4976 GcTGTTAACCTTCTGcCTTAATATAGAAAcCCCAAGCCCAAGCCcAGAcGACAGTAC 5035
| | | | |
QY 4560 accctagaagacttgcgaagaagataaaagaatagaaccagttctcgaacatccgaagg 4619
| | | | |
Db 5036 ACCcTTGAAGAcACTGGCAAGAcATTAATAAGATAGAcCAcTCTTGAGAcTCCGAGGGG 5095
| | | | |
QY 4620 accctgatacctcataatggaaggaatcctgcgcccaaaaagaggtttagaatalgtc 4679
| | | | |
Db 5096 AcCTGTATAcCTCATATATGGAAGAAATCTGCCCCAAAGAGGGGTATGAATATGTC 5155
| | | | |
QY 4680 caacagatcatcgtctaaacccaactaggaactaaacacctgcagcagttgtgaaga 4739
| | | | |
Db 5156 CAACAGATATcATGCTCTCAcCCAcCTAGGAcTAAAcCTGcACAcTGTGcTGAACA 5215
| | | | |
QY 4740 tccccctatcatgttctgaggttacagagatgtgtgactgcgtgtgtcaaacattgtg 4799
| | | | |
Db 5216 TCCcCTATAcATGTTCTAGAGcTACCAGAGAGTGGcTGAcTcGTGTCAAAcATTGTGTG 5275
| | | | |
QY 4800 cccctgcagactgtgttaatgttaatcccttcagaatacctccaagaaagagataaggga 4859
| | | | |
Db 5276 CCcTGcAGAcTGTGTTAAATGCTATATCTTCAGAGTAcCTCCAGAAAGAcATAAGGGGA 5335
| | | | |
QY 4860 agcccccagggcgtcactgcagggaagtgacttactaaggtgaagccggcttaataagga 4919
| | | | |
Db 5336 AcCCAcCCAcGGcGCTCAcTGGGAAGTGGAcTCAcTGAAGGTAAAGcCCGcTTAAATACGA 5395
| | | | |
QY 4920 aacaaatatctatgtgttltgttagacaccttltcaagatgtgtgtagagcttatcctact 4979
| | | | |
Db 5396 AAcAAATATCTATGTTGGTTTGTAGAcACcTTTCAGAGATGGTGAAGcCTTATTCCTACT 5455
| | | | |
QY 4980 aagaaagagacttcaaccgltgtgtgaagaataactcgaaggaatttccaaagatt 5039
| | | | |
Db 5456 AAGAAAGAcACTTCAACCGTGTGTAAAAAAATAcTGAAGGAATTTTCCGAGATT 5515
| | | | |
QY 5040 ggaatacctaaggttaatgaggtgcagacatgtgcagtttcgttggcccaagtagtag 5099
| | | | |
Db 5516 GGAATAcCTAAcGTATATGGGcTCAcAGAcATGGcTCAcGCTTTGTGcCCAGGTAAGTAG 5575
| | | | |
QY 5100 ggaactggccaagataatggggaatltgataaactgtcatgtgcatlacagacccaagc 5159
| | | | |
Db 5576 GGAcTGGCCAGATATTTGGGGATTGATTGGAACGTGCATTGTGCATAcAGAcCCCAAGC 5635
| | | | |
QY 5160 tcaagaacaggtagagagatgaatagaacatlaaagagaccccttaaccaattgaacca 5219
| | | | |
Db 5636 TCAGAcAGAGTAAAGAGATATTAACCAATTAAAGAGAcCCCTTAcCAAAATTGACACA 5695
| | | | |
QY 5220 ggaactggcatlaatgaatgtgatgtctcctgccttltgtctttaaaggtgaagaac 5279
| | | | |
Db 5696 GAGAcTGGcATTAATGATGTGATAGcTCTCTGcCCCTTGTGTGCTTTTAAGGTTAGGAGAc 5755
| | | | |
QY 5280 acccctgacagatttggcctgacccctatgaatgtctctacggggagacccccccttg 5339
| | | | |
Db 5756 AcCCcTGAGAcGTTTGGcCTGAcCCcCTTAATTAATGTCTAcGGGGAGAcCCcCCCGTGG 5815
| | | | |
QY 5340 gcaagaattgccttgcacataagtgctgatagtgtccttcccaagccttgttcttaagg 5399
| | | | |
Db 5816 GTAGAAATtGCTTCTGTATAcATAGTGTGATGTGCTGCTTCCAGCCCTGTCTTAAGG 5875
| | | | |
QY 5400 ctcaagggcctcgaatgtgtgagcagcagcgttggagaagcgtccggggagggcctatac 5459
| | | | |
Db 5876 CTCAAGGGcGCTCGAGTGGGTGAGGCACGAcCTGTGAAGAcGcTCCGGAGGcCTACTCA 5935
| | | | |
QY 5460 gg---aggaagacttgcaagttccacatgcttccaagtgtgagattcagttcatgttaga 5516
| | | | |

Db 5936 GGAGAGAGACTTGCAGATCCACATCGCTCCAACTGGAGATTCAGTCAATGTTAGA 5995
Qy 5517 cgcacccgtgaggaacacccctgagacacccgtgaggaagacattatcccttacttggac 5576
Db 5996 GCGCACCGTGGAGAAACCTCGAGACTCGGTGGAGGGCCCTTATCTCTGACTTTTGGAC 6055
Qy 5577 acaccaacgagctgtgaagtcgaagaatcccccacttgaatccatgattcccaagtaag 5636
Db 6056 ACACCAAGGCTGTGAAGTGGAGGAATCTCCACCTGGATTCATGATCCCAAGTTAAG 6115
Qy 5637 ccgagcgacacccctccgagctcggtgtgagagccgaaaaagactgagaatcccttaagct 5696
Db 6116 CTGGCGGACCTCCCGACTGGGGTGGAGGCCAAAGACTGAGAAATCCCTTAAACCTT 6175
Qy 5697 cgcctcatcgccgtgttcccttacttcaacaataaccccaagcgagtagtaaacgct 5756
Db 6176 CGCCTCATCGCCTGGTGTCTTCTTACTTCAACATTAACCTCCAGGCGAGTAAACGCT 6235
Qy 5757 tatagaagctcgaaaccccaatagaaccttacccttacttggctgattatgacccga 5816
Db 6236 TATAGACAGCTCGAACCCTCATATACCTTATCCCTTACCTGGCTGATTTATGACCTGA 6295
Qy 5817 taaggtgtcaactgtaataatgcaactcgaggtgttgcctccagaagcgacccgtgtgctga 5876
Db 6296 TACGGGTGTCACGTAAATAGCACTCGAGGTGTGCTCTAGAGGCACTGGTGGCTGA 6355
Qy 5877 actgacttctgcctccgattgattaaaccccgctgttaaaagacacacccccaactagt 5936
Db 6356 ACTGCATTTCTGCTCGGATTTGATTAAACCCGCTGTTAAAGCACACCTCCCAACCTAGT 6415
Qy 5937 ccgtagtataaggttctatgtctgcccaggaacagagaagaagaactgtggtgggttc 5996
Db 6416 CCGTAGTTATGGGTTCTATTGTGCTGCCAGGACAGAGAAAGAAATACGTGGGGGTTTC 6475
Qy 5997 ttgggaatccctctgtatgaggaatgagatgctgcacccccaagatggaacatggaatg 6056
Db 6476 TGAAGAAATCCTTCTGTAGGAGATGAGAGCTGCTGCTCAACGATGGAAGCTGGAATG 6335
Qy 6057 gccatctctctccaggaacgggtlaaaatctccttctgtaattccggcccgagcaagta 6116
Db 6536 GCCATCTCTCTCCAGGACCGGGTAAATTTCTCTTGTCAATTCGGCCCGGCAAGTA 6595
Qy 6117 caaaatgataataataataagatagagctgtcccccctcagacttgaattactataa 6176
Db 6596 CAAAGTATGAACTATATATAAGATTAAGAGCTGCTCCCACTGACACTTAATTAATGTAA 6655
Qy 6177 gataagttcaactg--aaagaaacagagaataattcaaaagtga laaagtgtatga 6234
Db 6656 GATAAGTTTCACTGAATAAGAAACGAGAAATATTCAAAAGTGAATAATGTTAGAG 6715
Qy 6235 ctggggaatagttttat--talaagcggtggagagcaggttccatttaacatctgcct 6293
Db 6716 CTGGGGAATAGTTTATTAATAATATGCGGGGAGGAGGAGTTCACATTAACTTGGCT 6775
Qy 6294 taggaatgaaagcgaggaacaaacccctgtgcaatgagaccccgataaagtaactgtcta 6553
Db 6776 TAGGATAGAGACGGGACAAACCCCTGTGGAGTGGAGCCGATTAATAATGTGCTGA 6635
Qy 6354 acagggcccccgcccttgagacacccga laaacttgccgtgtcccaatlaacccgct 6413
Db 6836 ACAGGGCCCCCGGCTTGAGAGCAACCGCATTAATTGCGGTCGCCCAATTAACCTGCT 6695
Qy 6414 ggcgctgacataacacagccgcttagcaacagtaacactgtatgatttccatacaaac 6473
Db 6896 GCGGCTTGACATACACAGCGGCTTAGCAACGATACACGATGATTCTTACCAACAC 6555
Qy 6474 ggcctagaactccccaggttctctgttgaagacagacagagacttgaatttccatca 6533
Db 6556 GCGTAGAACTCCCAAGTGTCTGTGTTAAGACAGGACAGACTCTTCAAGTCTCAATCA 7015
Qy 6534 ggagacttccagacataactcaacgagacccgtgacttctctgttgccttg 6593
Db 7016 GGGAGCTTTCAGGCCATCACTCCACGACCCGTATGCGACTTCTGTGGCTTGG 7075
Qy 6594 tctatccagagccctcccttataatgaagggaatggaatgaagaagaatlaactgtac 6653
Db 7076 TGTATCTCAAGGCTCCTTATTTATGAGGGATGGCTTAAGAAAGAAATTAATGTGAC 7135
Qy 6654 caaagagcatalaataatgataatgaggtgtccgnaaataagcttaccctcaatgaat 6713
Db 7136 CAAGAGACTAGAAATCAATGTATGAGGGGTCCGAAATAAGCTTACCTCAGTGAAGT 7195
Qy 6714 ttccgggaaggagactgatataggaagaagctccccaatcccaacaccccttctatag 6773
Db 7196 TTCCGGGAAGGGACATGATAGGAAGAACTCCCATTCACCAACACTTGTGTATAG 7255
Qy 6774 tactgtgttataagcaagcctcagaaatagattatagactgtgtlaaagaagt 6833
Db 7256 TACTGTGTTTATGACAGGCTTCAGAAATAGATTATAGTACCTGGTTATACAGGTG 7315
Qy 6834 gtgggcatgcaatgactgaggtttaaaccctgtgttccactgaagcttcccaacaa 6893
Db 7316 GTGGGATGCAATACAGGCTTAACCCCTGTGTTCCTCAGTCTTCAACCAATCCAA 7375
Qy 6894 agattctgttcatagttccaaatgctcccgagttgacttcaactccttgaaggtgt 6953
Db 7376 AGATTTCTGTATGATGTTCAATGTCCTCCGAGTGTACTACATCTGAGGAAGTGT 7435
Qy 6954 ccttgaatgataatgataatcgttataacccgaacaaagaagaccgatacccttaacct 7013
Db 7436 CCGTATGATATGACTATACGCTATACCGACCAAAAAGAGAACCGATTCCTTACCT 7495
Qy 7014 agctgtatgtcgtgatttgaaggagcggtgtggttgaagaacagagacagctgtccgtat 7073
Db 7496 AGCTGTAAATGCTGGAATTAAGGAGCGCGGTGGCTAGGAAGACAGGACAGCTGCTAT 7555
Qy 7074 caagagacacagcagctlagagaagaagactgtgtgacttcaatgtgcctgaagaaga 7133
Db 7556 CACAGGACACAGCACTGAGAAAGAGACTTGTGACTACTGCGCCATGACAGAAAGA 7615
Qy 7134 tctccagacctaagaaggagctgtttagaacacctaagaagctcctgacttctgtctga 7193
Db 7616 TCTCCAGGCTTATGAGAGGATCTGTATGACCACTAAGAGGCCGACTTCTTGTGCA 7675
Qy 7194 agtgtgttcaagaaacggagaggttagatgtgttcttcaagaaggttgggttg 7253
Db 7676 AGTGTCTTACAGAACCGGAGGGGATTAATGATGTGTTTCTAAGAGAAAGGTGATWG 7735
Qy 7254 tgcagccttaagaagaagtgttcttctatgtatgatacctcgaagacatcaagaactc 7313
Db 7736 TGCAGCCTTAAAGAAAGAAATGTTGCTTCTATGATACATCAGGAGCATCAGAGATC 7795
Qy 7314 catgaacagcttagaanaaagttagagaagcgtcgaagggaagaagagctgacagg 7373
Db 7796 CATGAGCAAGCTTATGAGAAAGGTTAGAGAGGCTGGAAGGGAAGAGGCTGACCAAGG 7855
Qy 7374 gtggtttgaaggatgttcaacaggttctcctgtgtgagacacccgtcttctgtctac 7433
Db 7856 GTGGTTTAAGGATGTTCAACAGGCTTCTTGGATGACACCCGCTTCTGCTGAC 7915
Qy 7434 ggggccccctagtagtccctgtcttcttacttaacagtgtggccttgaatlaagtt 7493
Db 7916 GGGACCCCTAGTAGTCTGCTGCTCTGTACTTATCAATGTTGGCCTTAAATTAATAGTT 7975
Qy 7494 tgttgccttlttgaagaagagtagtgcagttccagatcatgtaactgaagcaagta 7553
Db 7976 TGTTCCTTGTGTTAGAGAAAGGATGAGTGCAGTCCAGATCAATGACTTAAAGCAAGTA 8035
Qy 7554 ccaagagccttgaagcaagaagaaactgagccttccagcttccagttctaagttaga 7613
Db 8036 CCAAGGCTTCTGAGCCCAAGAGAACTGACTCTGAGCTTCCCAAGTTCAAGATTAGAA 8095
Qy 7614 ctatttaacagaagaaggttgggaatgaagaatgaagaatgaagaatccctccaga 7673
Db 8096 CTATTACAGACAGAGAGTGGGAAATGAAGGATGAAGGATGAAGGATGAAGGATGAAG 8155

```

- QY 7674 acccgaagtgtaataaagctctaaatgccccgaattacagaccctgtgctgcga 7733
      primer_bind
      /note="cap site"
      710..727
      /note="g1y4"
      1154..725
      /gene="gag"
      /gene="gag"
      1154..725
      /gene="gag"
      /function="retroviral capsid"
      /product="Gag protein"
      join(1154..725,2729..6313,6189..8162)
      /note="viral polypeptide"
      /protein_id="CAB65339.1"
      /db_xref="GI:6688946"
      /protein_start=1
      /db_xref="SPRMBL:O901x5"
      /translation="MGQVTVPLSLTLDHWTYRSRAHNLVSVYKKGWQFCASEMP
      TFDVGPSEGTNSETILLAVKAIIFQYTGSSHPDEPYLLTQDIAEDPPYKAPMLN
      KRKQPRILALGKNSHAERVSERPIYPIEPEPPYPPYPPYPPYPPYPPYPPYPPY
      PSAPGAPVVEGPAAGTSGRSGATEPDEIAILPLRTYGPMPGOLQLOLYPFSS
      ADLYMKTNHPPESEDPORLTGLVESISMHOPTMDCCOLLOTLFTTERERTILLA
      RKNVPGDGRPNOLONEIDMGPRPLRPGMDVYTAGRESIKITRQALVAGLRGASRP
      TNLAKREVMQGNPEPSEYLERIMARREPTDPPTSAQASVALAFITGQSLDIR
      KLIQRLGLOEALDLVNEAKVYRRETEREKQREKERERERDRQENLT
      KILAAVEEKSSEERERDRKIRSGPROSGINGRTPDKQCAVCKEGRHARNCPK
      KNGKPKVLTLEEDKDRGSDPLEPRTYLLVEGQPEYFLVDGAEHSLVLOPIGLK
      KKKSWMGATGORYPMTTRTDLAVGRVTHSLVPECPVPLIGDILLTKMGAOI
      SPEQGEVSVNNKPIYVLTLOLDEYRLSPQKPPDODIOSMLDPOFQANAEATGK
      LAKQVPOVIOJAKASITPVSAQVYPLSRARRGITPHVORLQOGLVLPVSPWMTPL
      LVRKRGINDYRPVODLRKVNKRVODIHTYVNPYNLSALPERNKWTVLTKDAKF
      CILRHPTSQPLPFWMDRGRTQOLMTPLPOFKNSPTTFEDBALHDLANERIOH
      POVTLQYVDDLALAGATKODLEGTAKALLLESLDGRASAKKQOIRREVTYIGS
      LRGQORMLTEARKKTIVOIPAPTTAKOYBEFGTAPFCRLMPGATLAAPLYPLKE
      KGEFVAPHOXKAPDAIKKALISAPALAPDTPKPTLYDERKVASGVITQTLGPM
      RBPVAVLSKIDPVASMPVCLAKIAVAAILVKDADKITLNGONTIVIAHLENVRO
      PPDKNMTNRMHTHOSLITERVTFAPRALPAPLLEDEDEPTHCOLDLEETG
      VKRDLDLPLTGEVLEWFTDSSYVEEKRMAGAVVGTHTMASLPEGTSAQKAE
      LMAILOALRLAIGKSIINITYDSRYAFATAHGAIVYKORGLTSGREIKKKEELST
      LEALHLPKRLAIHCPGHOKAKDLISRGNOMADRYAKOAOAVNLPLIIEPKAPEP
      ROYLEDMEQEKIKIDOFSEPTGTCYTSYGEKLLPKRGLGYVOOIHRLTHLGRKHQ
      OLVRTSPYHVLRLPGVADSPVSHCPQOLVYANNSRITPPGRKLSHGAIBEDPTE
      VAPKAYKXLYLVFDITSGWVEAPLPIKKESTVYAKLLEIEPREFIPVVISDNG
      PAEVAOSOGAKLIGIDMKLHCAYRPOSSQOYEMAKTIEITLTKLETGINDMIA
      LLPEVLFRNTPGQFGLTPEYELAYRPOSSQOYEMAKTIEITLTKLETGINDMIA
      LPEVLFRNTPGQFGLTPEYELAYRPOSSQOYEMAKTIEITLTKLETGINDMIA
      FORANKOLREAVSGEBDOVPHRFOVGVSYVRRRAGNLETFRKGPVLLTTPAV
      KYEGISTWHTSHVRLAPDPSGWAETEMDKLRLRLIYVYNNNSPOGIIHTPLSH
      RULPRGCEPKRLRPLISFASITAMPLITITTPQASSKLIDSSNPHRLSLITMLIDP
      DTGVTVNSTRGVAPRGITWPELHFCRLRLINAVASTPENVLRSGFYCCPETEREKC
      GDGSEFCRMSVCVINDGDKMWPISLQDRVKFSVNSGPKRYVMKLYKXKSCSPSD
      LDYLKISTETEGKOENIKWINGMSGIVFYKYGGAGSTLIRLLETGTEPPVAVG
      FDKVLAEQGPALPEPHNLVYPOLTSLRCDLTOSPSNGTGLITPMPRNSGPVAT
      GORLESLIOGAFOALNSTDPATSCMLSSGPPYVGMKKEKFNVRKRNQCTW
      GSRNKLITLTVSGKGTGCKAPRSHOHCISTYVYEOASMOYIYVPGYNNRMAKNTGL
      TPCVSTSVFNOSKDFCWQVQIYPRVYTHPEEVVLEDEVDYRNRKREVSITLAVMG
      LGTAVGVGTGTAALITGPQOLEKGLGELHAAMATDEYRNRKREVSITLAVMG
      ONRGLDILFLREGICLALKECCFCFYVDHGAIRDSKFLRELRERRREREDOG
      FEGMFRSPMMPTTLLISALGTPLVILLITLVGPCILNFFAVHREVSANGIMVLRQO
      YGILSGEETDL"
      2729..6313
      /gene="pol"
      /function="retroviral protease and polymerase"
      /product="pol protein"
      2729..6313
      /gene="pol"
      /gene="pol"
      6189..8162
      /gene="pol"
      /function="env"
      /product="env protein"
      6189..8162
      /gene="env"
      /gene="env"
      8212..8918
      /note="grown in PK15 (ATCC No. CCL-33) cell line"
      1..707
      LTR misc_feature 545
      LTR misc_feature 545

```


QY 2100 aactgcccaagaaggaaacaagaacgaagtccttaagcttagaagaagataaagat 2159
DB 2666 AACTGCCCAAGAAGGGAACAAGGACCGAAGCTCTTAAGATAAGAAAGAT 2725
QY 2160 taaggaggaagggttcggaacccctcccgagcccggttaacttgaagtggaggg 2219
DB 2726 TAGGGGAGCGGGGTGGACCCCTCCGAGCCAGGGTAACCTTGAAGGTGGAGGGG 2785
QY 2220 caacgaattgagttcctgattgataccgagcgagacatcagtgctgctacaacatla 2279
DB 2786 CAACCAAGTGAAGTTCCTGTTGATACCGAGACGAGCATTCAGTGTGCTCAACACCTTA 2845
QY 2280 ggaactaaagaaaaaaalcctgggtatgggtgccaagggcaagcgagtalcca 2339
DB 2846 GGAANAATTAAGAAAAAATCCTGGGTATGGGTGCCACAGGCAAGCGATATCCA 2905
QY 2340 tggatcccggaagaaccgttgattggagtggaagggtaaccacactgcttctgctc 2399
DB 2906 TGGACTACCGGAAGAACCGTTGACTTGGAGTGGAGCGGGTAAACCACTCGTTCTGCTC 2965
QY 2400 atccctagtgccagtaaccctcttagtgagagacttaagccaagatggagatcaa 2459
DB 2966 ATCCCTAGTGCCAGTACCCCTTCTAGTAGAGACTTACTGACCAAGATGGAGCTCAA 3025
QY 2460 attcctttgaacaagaagaccagaagtgctgtgataaacaacccatcaactgtgtg 2519
DB 3026 ATTTCTTTGAACAAGGAACGAGAAAGTGTGTGTAATTAACAACCCATCATCTGTGTTG 3085
QY 2520 accctcaattatagatgataatgactataltccccaaglaaagcctgaatcaagat 2579
DB 3086 ACCCTCAATTAAGATGATGATATCACTATATCTCCCAAGTAAAGCCTGATCAAGAT 3145
QY 2580 atcaagtcctggttgaggaagcttccccaagcctggagcaaaacgagagatgggttg 2639
DB 3146 ATACAGCTCGGTTGGAGAGCTTTCCCAAGCCTTGAGGCAACCCAGGAGTGGGTTTG 3205
QY 2640 gcaaaagaaagtcctcccaaggttaltcaactgaagcgagtgctacacagatcagtc 2699
DB 3206 GCAAAAGAAAGTCCCGACAGGTTATTAACCTGAAGGCCAGTGTCTACACAGTATCAGTC 3265
QY 2700 agacagtaacccttgatgagagaggtcgaagaagaaattggcgcaatgtccaagata 2759
DB 3266 AGACAGTACCCCTTGAGTAAGAGGCTCGAAGGAAATTTGGCCGATGTTCAAAAGATTA 3325
QY 2760 atccaaagggcatcctcagttcctgctcaatcccttggaatactccctgctacaggt 2819
DB 3326 ATCCAAAGGGCATCTCTAGTCTGTGCCAATCCCTTGGAATATCTCCCTGTACCGGTT 3385
QY 2820 aggaagccttggagccaatgatalcgaaccagtaacagacttgaagaggltcaataaag 2879
DB 3386 AGGAAGCCTGGAGCAANTGATATGACCAAGTACAGGACTTGAAGAGTCAATATAAGG 3445
QY 2880 gtgcagagataatacccaaggtcccgaaacctataacctctgaagcgctcccgct 2939
DB 3446 GTGCAGAGATATACCAACCAAGGTCGGAACCCCTTAATACCTGTGAGCCCTCCGCTC 3505
QY 2940 gaaagaaactgtaacagataltggaactaaagatgcttcttcgctgagatlaaac 2999
DB 3506 GAAGGAAGTGTGACACCACTATTGGACTTAAGAGATGCTTCTTCTGCTTGAGATTAC 3565
QY 3000 cccactagccaacacttttgccttcgaatgagagatccagtgtaacggaagaaacgg 3059
DB 3566 CCCACTAGCCAAACCGCTTTTGGCTTGAATGAGATCCAGATCCAGGAAGAACCGGG 3625
QY 3060 cagctacttgagcccgactgcgccaaggttcaagaactcccgacaactcttgaagaa 3119
DB 3626 CAGCTACTTGACCCGACTGCCCCAAGGGTTCAGAAATCCCCACACATCTTTGACGAA 3685
QY 3120 gacctacaggaagccttgccaactcagatcccaacccctcaggttaacccctcctcag 3179
DB 3686 GCGCTACAGAGGAGACTGGCCAACTTCAAGATCCAAACCCCTCAGGTGACCTCTCCAG 3745
QY 3180 taagtgatgacctgtcttgccggagagccaacaaagactgcttagaagtgtaagaa 3239
DB 3746 TACGTGATGAGACTGCTTGTGGGGGAGGCCACCAACACAGACTGCTTAAGAGTACGAAG 3805
QY 3240 gcaactgctgtaactgtctgaactaggtcacaagcctctgtctaaagaagccagat 3299
DB 3806 GCACCTACTGTAATTTGTGTACCTAGGCTACAGAGCTCTGTGTAAGAAAGCCAGATT 3865
QY 3300 tgcagagaagaggttaactacttgggtgacgttttgccgggacggagatgagctgagc 3359
DB 3866 TGCAGAGAGAGGTAACTACTTGGGGTACAGTTTGGGGGGGACGAGATGGCTGAGC 3925
QY 3360 gaagcaggaagaanaactgtaagcagataccggcccccaacacagccaagaatgaga 3419
DB 3926 GAGCGACGGAAGAAACTTATGTCAGATACCGGGCCCCACACAGCAACCAAGTGA 3985
QY 3420 gaagtttgggaacagctggaatttgcagactgtgataccggggttttgcagacttaca 3479
DB 3986 GAGTTTGTGGGACAGCTGATTTTGCAGACTGTGGATCCGGGGTTTGGCACTTAGCA 4045
QY 3480 gccccaactlacccgttaacaaagaaggggattctcctggtctcctgaagcagcag 3539
DB 4046 GCCCACTCTACCCGCTTAACCAAGAAAGGGGAATTCCTCGGGCTCTGAGCACGAG 4105
QY 3540 aaggaatttgatgatacaaaaaggccctgtagcgcaacctgctctgagccctcctgac 3599
DB 4106 AAGGCAATTTGATGATATCAAAAAGCCCTGCTGAGCGCACCTGCTCTG6CCTCCCTGAC 4165
QY 3600 gtaactaaaccccttaaccccttaagtgtgataagcgtlaaaggagtagcccgaggttla 3659
DB 4166 GTAACATAAACCTTTACCTTTATGTGATGAGTGAAGGGAGTGAAGCCGAGGATTTTA 4225
QY 3660 acccaaaccccttagagcaatggagagacgtgttgcctacgttcaagaagcttgaacct 3719
DB 4226 ACCCAACCCCTTAGGACATGAGAGGAGACACTGTTGCTACCTGTCAAAAAGCTTGATCT 4285
QY 3720 gtagcagtggttggcccgatgctgtaaggtatcgcgaagctgtgagccatagtgtaag 3779
DB 4286 GTAGCCAGTGGTGGCCCGTATGCTGTGAAGGTATGCAAGCTGTGGCCATCTGCTCAAG 4345
QY 3780 gaagctgaacaaatgacttgggacagaataaactgtataagcccccatgcatgtgag 3839
DB 4346 GAGGCTGACAAATTAATGACTTTGGAGCAGATTAATTAATGAGCCCCAGCATTTGAG 4405
QY 3840 aacatcgttgcgcaagccccagacagatgtagtaacaaagcccgatgaagccccatcaa 3899
DB 4406 AACATCGTTGGCAACCCCGACGAGATGATGACCAAGCCCGATGACCCACTATCA 4465
QY 3900 agcctgcttcaagagaggttaacttgcctccacagcgctctaaaccccgcaact 3959
DB 4466 AGCCTGCTTCTACAGAGAGGGTCACTTCCGCTCCACAGCCGCTTCAACCCGTGCCACT 4525
QY 3960 ctctgctgaagagactgtaaacagtgatcatgatalgcatacaactatgtatgag 4019
DB 4526 CTCTGCTGAAGAGACTGATGAGACCACTGATGATGTCATCAACCTTATTTAG 4585
QY 4020 gagactgggtccggaagaccttaacagataacccgtgactgtagaaggtgtaacctgg 4079
DB 4586 GAGACTGGGGTCCGAAAGACCTTAACAGACTTACCGCTGAGAGAGGTCTAACTGG 4645
QY 4080 ttaactgaaggaagcagatgltgtgaaagtgaaagatggtgtggcgaggtgtg 4139
DB 4646 TTCACTGAGGAGACAGCTATGTGTGAAGGTAAAGAGATGCTGGGGCGCGGTGTG 4705
QY 4140 gaeggagccagacagatctgggcagcagccttgcgggaaggaactcagcgcaaaagct 4199
DB 4706 GAGCGACCCACAGATCTGGGACAGCACTGCGGCAAGGAATCTGAGCGCAAAAGCT 4765
QY 4200 gagctatgagccctcaagcaagcttgcggtgtgcggaaggaattccataaactat 4259
DB 4766 GAGCTCATGGCCCTACAGCAAGCTTGTGGCTGGCCCAAGGGAATCATTAACATTTAT 4825
QY 4260 acgagacagagtgatgcttctgacgtgcaacgtacaggggacatctaaagaagg 4319

Db 4826 ACGGACAGCAGGTATGCTTTGGGACATGCACACGTACACGGGGCCATCTATATAACAAAGG 4885
Qy 4320 ggggtgttacctcaagcaggaggaataaagaacaagaagaaattcaagctatla 4379
Db 4886 GGGTGGCTTACCTCAGAGGAGGGAATATAACAAAGAGAAATTTCTAACCTATTA 4945
Qy 4380 gaagccttaccttgcocaaagagctagctattatacaactgtccctgacatcagaagacc 4439
Db 4946 GAAGCCCTTACATTTGGCAAAAGGCTAGCTATTTATACACTGTCTGTGACATCAGAAAGCC 5005
Qy 4440 aaagatctcatctatagaagaaaccagatgctgacccgggttgcagaagcagaccag 4499
Db 5006 AAAGATCTCATATCTAAGAGGAAACCATATGCTGACCGGTTTGCCTAAGAGGCGCCAG 5065
Qy 4500 gctgttaacctctgacctataatagaacacgcccagaagccccaagaacacagacagtlac 4559
Db 5066 GCTGTAACTTCTGGCTATATATAAAGCCCAAGGCCCAAGACCAGACAGCAGTAC 5125
Qy 4560 accctagaagaactggcagaagataaagaatagaccagttctctgagactccggaagg 4619
Db 5126 ACCCTAAGACACTGGCAAGAGATATAAAGATAGACACAGTTCCTAGACTCCGAGAGGG 5185
Qy 4620 accgtcatacctataltggaagaaatctgcccacaagaagaggttagaataatgc 4679
Db 5186 ACCGTATATACCTCATATGGGAAGAAATCTGCCCCACAAAGAGGTTAGAAATATGTC 5245
Qy 4680 caacagatatcogtctaaccccaactagaagactaaacaactcagcagatgtgtcagaaca 4739
Db 5246 CAACAGATATACCTGCTCACCCACCTAGGAATTAACACCTGCAGAGTGTGTCAGAACCA 5305
Qy 4740 tccctcatcatctgttcgaaggtcacaagaggtggtgactcaggtgttcaaaacttctg 4799
Db 5306 TCCCTTATATATGTTCTGAGAGCTACAGAGAGTGTGACTGGTGTGCTCAAACTTGTGTG 5365
Qy 4800 cccctgcaagcttggltlaatglaactccttccagaatacctccagaagaagactcaaggga 4859
Db 5366 CCGTGCAAGCTGGTTATGCTAATCTTCCAGAAATACCTCCAGAAAGAGACTAAGGGGA 5425
Qy 4860 agccaccccaagcgtcactcbyggaagtggacttcaactgaggttaagccggtcctaatacga 4919
Db 5426 AGCCACCCAGGCGCTCACTGGGAAGTGACCTACTGAGTAAAGCCGGCTTAATTCAGGA 5485
Qy 4920 aacaatcatctattgtgttctgtagaacacttccagagtggttagaggttactact 4979
Db 5486 AACCAATATCTATTGGTTTGTGTAGACACCTTTCAAGATGGGTAGAGGCTTATCTACT 5545
Qy 4980 aagaagaagactcaacccgtgtgtgtctaaagaanaatactggaaggaatttttccaagaatt 5039
Db 5546 AAGAAAGAGACTCAACCCGTGGTGGCTAATAAATACTGAGGAATTTTCCGAGATTT 5605
Qy 5040 ggaatacctaagtaataaggtcagacaatgtgtccagcttctgtgcccagaagtaagtcag 5099
Db 5606 GGATATACCTTAAGTAAGGTAACCGGGTCAAGCAATGTCTCAAGCTTTTGTCCCAAGTAACTCAG 5665
Qy 5100 ggaactggccaagatattggaggtattgtgaactcattgtgtcatcagacccccaagc 5159
Db 5666 GGACTGGCCAAAGATATTGGGGATTGATTTGAAACCTGCATTGTGCTATACACACCCCAAGG 5725
Qy 5160 tcaaggaagtagaagagatgaatagaacatlaaagagaccccttacaatltgacaca 5219
Db 5726 TCAGGACAGGTAGAGAGATGATAGAACATTAAAGAGACCCTTACCAATATGACCACA 5785
Qy 5220 gaagctggcaatlaagatgtgactcctcctgcttctgttctttagagtgagaac 5279
Db 5786 GAAGCTGGCATATATGATTTGATAGCTCTCCCTCTCTTGTGTGCTTTTAGGGTTAGGAAC 5845
Qy 5280 acccctggaacagtttggtgtcagcccccataatgtatctcagcggggaacccccccgttg 5339
Db 5846 ACCCCTGGACAGTTTGGGCTGACCCCTATGAAATGTCTCTACGGGGAGACCCCCCGTTG 5905
Qy 5340 gcaagaatltgcttgcacatagtgctgagatgtgtgtcttccagccttgttctctag 5399
Db 5906 GTAGAAATGTCTTGTGACATAGTGTGATGTCTGTCTTCCAGCCTCTGTCTCTAG 5965

Qy 5400 ctcaagcgcctcagtggtgtgagcagcagcgttgaagcagctcccggaagcctactca 5459
Db 5966 CTCAGGGCGCTGATGGGTGAGGCAACAGAGGTGGAAGCAGCTCCGGGAGGCCACTCA 6025
Qy 5460 gq---aggagacttgcaggttccacatcogcttccaagttgagatcagttcattgta 5516
Db 6026 GGAGAGAGGAGACTTCCAAAGTTCCACATGCTTCCAAAGTGGGAGATTCAGTATATGTA 6085
Qy 5517 ggcacccgtgaggaacactcagactcgttgaagaggaacctatctgacttctgacc 5576
Db 6086 CGCCACCGTGCAGGAACCTCGAGACTCGGTGGAAGGGCCCTTATCTGTACTTTGACC 6145
Qy 5577 acaccaagcgtctgaaagtgcgaaggaatcccaactggaatcattccacogtlaag 5636
Db 6146 ACACCAAGCGGCTGTAAGTGCAGAGAAATCTCAACCTGATTCATATCCACAGTTAAG 6205
Qy 5637 ccggtgcacactcccgactcgggtgtgagagcgcgaagaaacttggaatcccttaagctt 5686
Db 6206 CTGGGGCCACCTCCGACTCGGGGTGAGAGCCGAAAGACTGAGAAATCCCTTAAGCTT 6265
Qy 5697 cgcctccatcgcctgtgttccctactactaacaatacctccagcagtagtaaacgct 5756
Db 6266 CGCTTCATCGGCTGTTCTTACTTACTTAACAATACTCCAGCGCACTAGTAAAGGCT 6325
Qy 5757 tatagaacgttcgaacccccatagaacttlatcccttaacctgtgtattatgaacctga 5816
Db 6326 TATAGACAGCTTCGAACCCCATAGACCTTATCCCTTACCTGCGGATTTATGACCTGA 6385
Qy 5817 tacgggtgtcactgttaaatagacactcgaaggtgtgtctctatgaagcactgtgtgcctga 5876
Db 6386 TACGGGTGTACTGTAATAGACTCGAGGCTGAGGCTGCTCTTGAAGGCACTGTGGTCTGA 6445
Qy 5877 actgtacttctgctcgcgatttgatlaaaccccgctgttlaaagcacacctcccaactagt 5936
Db 6446 ACTGCATTTCTCTCCGATTGATTAACCCGCTGTAAAGCACACCTCCCAACCTAGT 6505
Qy 5937 ccgtatgtatgtgttctatctgtctgcgaagcagaagaagaagaatactgttgggtgtc 5996
Db 6506 CCGTAGTAAAGGGTCTTATGCTGCCAGGCAACAGAAAGAAATCTGTGGGGATTC 6555
Qy 5997 tgggaatcctctgttaagagaatgagctgtcaccctcaagaatgtagacttgaagt 6056
Db 6556 TGGGGAATCTCTGTAGAGATGAGAGCTGCTACCTCCCAACATGAGATGGAATG 6625
Qy 6057 gccgaatctctccagaacccggtlaaactctccttctgtcaacttcggccggccaagta 6116
Db 6626 GCCGATCTCTCCAGGACCGGGTAAATTCCTTTGTCAATTCCGGCCGGCAAGTA 6685
Qy 6117 caaatgtatgaactatataagaataagactgtctcccaatcagacttaagtattctaa 6176
Db 6686 CAAAGTGAAGAACTATATAAGATMAAGCTGCTCCCATCAGACTTATGATTA 6745
Qy 6177 gataagttcactg--aaaggaacaggaataatctcaaaagtgtatlaag 6234
Db 6746 GATTAAGTTCACTGAAAAAGGAAACAGGAAATTTTAAAGTGTGATTAATGCTATGAG 6805
Qy 6235 ctggggaatagttttat-tataatgcggggagcagaggtccacttaacatctgct 6293
Db 6806 CTGGGGAATAGTTTATTAATATGCGGGGAGACAGGTCACACTTAAACATTCGCT 6865
Qy 6294 taggataagagacgggaaacaaacccctgtgtgcaatggagaccgaataagtaactgtcga 6353
Db 6866 TAGGATAGAGAGCGGAGAGAAACCCCTGTGCTGCAATGGAGCCGATTAAGTACGCTGA 6925
Qy 6354 acaaggcccccggtcgtgagccacgcatcactgtcgggtgccccatlaactcgt 6413
Db 6926 ACAGGGGCCCGCGGCTGAGAGCCAGCATTAAGTGTGCGGTCGCCAATTAAGTCTGCT 6985
Qy 6414 gcggcctgacataacacacccgctagacacagtaacacatgattgattcctacacac 6473
Db 6986 CGGCGCTGACATTAACACAGTCCGCTACGACAGTAACTGATGATTTCTTACCAACAC 7045

QY 6474 gctagaactccccaggtgttcttgaagacagagacagacttcaatctatca 6533
 DB 7046 GCGTAGAACTCCCAAGGTTCTGTATAGACAGACAGACTTGTAGTCTATCCA 7105
 QY 6534 gggagcttcccaagcatcaactccacccagacctgtatgcacttcttctgtgttg 6593
 DB 7106 GGGAGCTTCCAGGCATCACTCCACGACCTGTATGACACTTCTGTGTGGCTTGG 7165
 QY 6594 tctatcccaaggcctccttaltatgaaggagatgctaaagaagaatcaatgtgac 6653
 DB 7166 TCTATCTCTCAGGCTCTCTATATAGAGGAGATGCTAAAGAAATTCATATGTGAC 7225
 QY 6654 caaagagcatagaatcatgtaacatgtgggtccgaaataagattacactcaatga 6713
 DB 7226 CAAGAGCATAGAAATCAATGATACATGGGGGTCCGAAATTAAGCTTACCTCACTGAAT 7285
 QY 6714 ttccgggaaggagacatgataagaaagctcccccatcccaacacacttctgtatag 6773
 DB 7286 TTCGGGGAAGGGACATGACATAGGAAAGCTCCCATCCACCAACACTTGTGTATAG 7345
 QY 6774 tactgtgttaltatgacagagcctcagaatacaatgattatgaactgttataacagtg 6833
 DB 7346 TACTGTGTATTATGACAGAGGCTCCAGAAATCAGTATTAATTAACAGGTG 7405
 QY 6834 gtggagatcaactactgggttaaaccctgtgttccacactgaacttcaacacacaa 6893
 DB 7406 GTGGGATGCAATACTGGGTTAACCCCTGTGTTCACCTCAGTCTTAACCAATCCAA 7465
 QY 6894 agattctgttcaatgtgtccaaatcgltcccgagtgtaactacactctgagaaatgtg 6953
 DB 7466 AGATTTCGTGTCTAGTGGTCAATGTCGCCCGAGTGTTACTACATCCTGAGAAATGGT 7525
 QY 6954 ccttgatgaataatgactatctgtatataacagcaaaaagaagaccgtatccttacc 7013
 DB 7526 CCTTGAAGATATGACTATGCTATTAACGACCAAAAGAAAGAACCCGTATCCCTTACCT 7585
 QY 7014 agctgaatgtcggatctgaagacgagcgttggcgttaggaacagagacgtcctgtat 7073
 DB 7586 AGCTGATATCTCGGATTTAGGAGACGGCGCTTGGCTTAGAGACAGGACAGCTCCCTGAT 7645
 QY 7074 caaagacacacagcagctagagaagaagacttgtgaactacatgagcagatgacagaaga 7133
 DB 7646 CACAGGACACACAGCAGCTAGAGAAGACTTGTGAGCTACATGCGCATGACAGAGAA 7705
 QY 7134 tctcccgagccttaagagagctgttagcaacctagaagaagctccttacttctgtctga 7193
 DB 7706 TCTCCGAGCCTTAGAGAGTCTGTAGCAACTTAGAAGAGTCCCTGACTTCTTGTCTGA 7765
 QY 7194 agtgttctacagaacccgagggagatagaactgtctgttcttaagaagaagtgtgtatg 7253
 DB 7766 AGTGTCTCTACAGAACCGGAGGGATTAGATCTGCTGTTCTTAAGAGAGTGGTTATG 7825
 QY 7254 tgcagccttaaaagaagaatgtgtcttctatgtagtaactcaagaagcaatcagagact 7313
 DB 7826 TGCAGCCTTAAAGAAAGAAATGTCTCTATGATGATCACTGAGGACATCAGAGACATC 7885
 QY 7314 catgaagaacactagaaaaaagttagagagcgtcgaagggaagaagaaggtctacagagg 7373
 DB 7886 CATGAGCAACCTTAGAAGAAAGTTAGAGAGGCTCGAAGGGAAGAGAGGCTGACCAAGG 7945
 QY 7374 gtgtgtgaagagatgtgtcaacaggtctccttgaatgacacaccttcttctgtctgac 7433
 DB 7946 GTGTGTGAAGAGATGTTCAACAGGTCTCTTGGATGACCACTCTGCTTCTGTCTGAC 8005
 QY 7434 ggggacctagtagtctgtctctgttacttacaagtggagccttcttcttaataatggt 7493
 DB 8006 GGGACCCCTAGTATCTCTGCTCTACTTACTTACAGTTGGGCTTGTCTTATATAGGT 8065
 QY 7494 ttgtgcttcttgaagaagagatgagtagccagataatgttactttaggaagaaga 7553
 DB 8066 TGTGTCTTGTGTTAGAGAAAGATGAGTGCAGTCCAGATATGTGTTAGGCAACAGTA 8125
 QY 7554 ccaagccttcttgagccaagagaagaactgaccttaccagcttccaaagattcaagata 7613

DB 8126 CCAAGGCTTCTGAGCCCAAGGAAACTGACCTTACGCTTCCAGTCTTAAGATAGAA 8185
 QY 7614 ctattacaagaagaagtgtgggaatgaagaatgaatgaactcaactccacaga 7673
 DB 8186 CTATTAAACAGCAAGAAGTGGGAAATGAAGATGAAGAAATGCAACCTTCCACAGA 8245
 QY 7674 acccaggaagttaataaagaactcaaatgccccgaattacaacctgtgtctgca 7733
 DB 8246 ACCCAGGAAGTTAAATAAAGCTTAATGCCCCGGAATTCAGACCTGCTGGCTGCA 8305
 QY 7734 gtaaatgtagaaggtcacacttccatttctccagagcctgtatccttgcctaagta 7793
 DB 8306 GTAAATAGTAGAAGTCACTTCTCTATGTTCCAGGGCTGCTATCTGGCTTAAGTA 8365
 QY 7794 agatacagaagaatgagttgactaatgccttactgtatctgtttaaactctgagacc 7853
 DB 8366 AGATTAACAGGAATGATGACTAATCCCTATCTGATGATGTAATAACTGACCTGGACC 8425
 QY 7854 atagaagaattgattacacattgacagccctagtgactatcctaactgaactgtcac 7913
 DB 8426 ATGAGAAGATTTGATTACACATTGACAGCCCTAGTGACCTTATCTCACTGCAATCTGTAC 8485
 QY 7914 tctgccagagagcccaagcagatgagacactccagagc-----tatt 7951
 DB 8486 TCTGCCAGAGACCCACGACAGATGCGGACCTCCGAGCTATTTAAATGATGGTCCAC 8545
 QY 7952 -----tatt 7955
 DB 8546 GGAGCGCGGCTCTCGATATTTTAAATGATTTGTCACGAGCGCGGCTCTCGATATTT 8605
 QY 7956 ttaaatatattgttccagagagcggcgtctcgtatattttaaagtattgttcaatgga 8015
 DB 8606 TTAATATGATTTGTTCCAGAGAGCGGGCTCTCGATATTTTAAATGATTTGTTCCACGGA 8665
 QY 8016 ggcgagctctcgtatattttaaaltgattgttltgtgacagcagagcttltgtgaacc 8075
 DB 8666 GCGCGGCTCTCGATATTTTAAATGATTTGTTTGAACGACAGCGCTTGTGGAAC 8725
 QY 8076 ccataaagcgttcccgatctccgacactggggcgagactctctacaccttgcgtgtgta 8135
 DB 8726 CCATTAAGACCTGTCCGATTCGCACTCGGGGCGGCAATCTCTACCTCTCGTGTGTA 8785
 QY 8136 cgactgtggcccgagcgcgtgtgaataaataatccttgcgtgttgcacaa 8191
 DB 8786 CCAGCTGTGGGCCCGGCGCTTGAATTAATCTCTTCTGTTTGCATCAGGA 8841

RESULT 6
 PERY17013 7808 bp DNA VRL 14-APR-2000
 LOCUS Porcine endogenous retrovirus gag, pol, env genes.
 DEFINITION Y17013
 VERSION Y17013.1 GI:3005985
 KEYWORDS env gene; gag gene; pol gene.
 SOURCE porcine endogenous retrovirus.
 ORGANISM Porcine endogenous retrovirus; Retroviridae; Mammalian type C retroviruses; 1-Mammalian type C virus group.
 REFERENCE 1 (bases 1 to 7808)
 Czudenza, F., Fischer, N., Boller, K., Kurth, R. and Tonjes, R. R. Establishment and characterization of molecular clones of porcine endogenous retroviruses replicating on human cells
 J. Virol. 74 (9), 4028-4038 (2000)
 JOURNAL MEDLINE 20219394
 REFERENCE 2 (bases 1 to 7808)
 Tonjes, R. R. Direct Submision
 JOURNAL Submitted (30-MAR-1998) R. R. Tonjes, Paul-Ehrlich Institut, Paul-Ehrlich Str. 51-59, P.O. Box 1740, D-63207 Langen, FRG
 FEATURES Location/Qualifiers
 1..7808
 /organism="porcine endogenous retrovirus"

/db_xref="taxon:61673"
/tissue_type="kidney"
/cell_line="PK15"
260..1834
/gene="gag"
260..1834
/gene="gag"
/codon_start=1
/protein_id="CA76581.1"
/db_xref="GI:3005986"
/db_xref="SPTREMBL:O73504"
/translation="MGQVTTPLSLTLDHMTVEVSRANHSVQVKGKPMWQFCASEMP
TFDVGWSEGTENSEILIAVKAIIQTGPSSHPQDEYIITFODLADPPVPMV
KPKRGKRLALGKKNHSAKEVPEPRITYPEIEEPDPEPPGPPPAOGAVG
PSAPGAVVGPAGTSGRRGATPERTDEIILPLRTYGPMPGQLOLPQVYPS
ADLYNMKNHPSEDPQRLGLVSEISMESHPQWDCOOLLOTFLAERERIL
RKNVPGADGRTQLONETDMGFLTRIGMDNTAEGESLKITQALVAGLRSR
TKLQREGLQAELELDLVREAKEYVGRETEEEKERERERERERERERER
KILAAVEGKSSRERDFRIRSGPQSGNLGNRPDLKDQCAVCKEKHMANRCK
KGNKPKVLALEDKD"

gene
CDS
2120..4726
/gene="pol"
2120..4726
/gene="pol"
/codon_start=1
/protein_id="CA76582.1"
/db_xref="GI:3005987"
/db_xref="SPTREMBL:O73505"
/translation="MGAQISFEGRPVSVNKPITVLTQLDLDEXRLYSPOVQPD
IOSWLEQFQAMETAGMLAKOVPOVQLKASAPVSVKQPLSREAREGIMPHO
RLIOGILVPOSWPNTPLVRKPGNDYRPODLREVNKRVODIHPVPMVNL
ALPERMWTIVLDKDAFECFLRLHPTSOPLFAEMRPGTGRQOLWTRPLRPF
PTIDEALHRDLANERHOPOVLTLOYVDLLAGATKODCLEGTALLLEISDL
ASAKKOICREVTYITSLRGQORLTERKKTVOIPAPTTAKOVRFLGTAGCR
LWIDFATLAPLYPLTKEGESMAPEHOKADAIKALLSAPALADPVTPTLY
VDERKGVARVLTQILGPMRPVAYLSKIDPVSAGPVLKAIAAVATIKRADLT
LGONITVIAHVALENTIROPDRMMTMAKMTIHVOSILTERVTEAPALNATIL
ETDEPVTCHDCHOLIETGVKRLDTIDPIGEVLTFTDSSIVYEGKRMAGAVD
THTTMASTPEGSAOKAEMALTOAIRLAEGKSINITYDSRYAFVAHGAIVQR
GLTASGREITKNEELISLEALHLPRKRLIHCPOHOKAKDIISNGMADVAKOA
AOAVNLPIETTERAPRPRQYLTLEMOEIKIDQSETEGTCYSYKELIPHEG
LEVVOOHLRLHGLTKHLOLVRTSPYHVLRLPGVADSVVKKHCPOLVANSRIL
GKRLRSGHPGAWHEVDTEVPAKYGNKLYLVVDFFSGWEAVYPRKETSIVAKY
WRKFFRLELYR"

gene
CDS
5294..7267
/gene="env"
5294..7267
/gene="env"
/codon_start=1
/protein_id="CA76583.1"
/db_xref="GI:3005988"
/db_xref="SPTREMBL:O73506"
/translation="MHPILSMRLPTRGGEPRRLRILPSFASIAFWLTITLTPQAS
RLIDSSNPHRPLSLWLIDPDTGVNSTRGVAPRTGWPMLHLCRLINAVASTP
PMLVRSYGYCCGTEKEKYGCGSESEPSRMSCTNSNDGMKMLISIDRPFSEFN
SGGKYKVMKLYKDKSCPSDLVYLIKISTTEKQKQENLOKMGNSKGVIFKYGGA
GTTTLRLRIETGCTEPRAVAPDVKYLADEGPPALPEPHLVPOLTSRPLDTOP
GTTGLIPTTPRNSPEVPKTSQRLSLIOLGAFQAINSDIPATSSCMCLSGSPY
EGMAKGRKNTKREHNOCTGWSRNLKLTLEVGKGTGICGAPPSHOHLCTSYVEO
ASENDYLVGVNMAACNTGLTVCSSSVFNOSKPCVAVQIVPVVYHPEVVLDEY
DYRNNRKEPVSLLAVMLGTAVGSTGTAALITGPQOLEKIGLGHAAMTDLR
ALBESYNSLESLTSLSEVYONRGLDILPLREGGLCAALKEBCCFFVDHGAIRDS
MSKLRRLRRRRRRRBDAGWEGFNMNSPMTTLLSALTGLVLLLVGLLVGCLIN
RFAVAFREVGAVQIMVLRKQYQGLLSOGETDL"

BASE COUNT 2077 a 1971 c 2047 g 1713 t
ORIGIN

Query Match 93.1%; Score 7644; DB 14; Length 7808;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 7746; Conservative 0; Mismatches 55; Indels 9; Gaps 6;

QY 329 atcagcagcagctgctagtagagatatacagcgtgcaccccttgaggagacgcccggaggtg 388
Db 1 ATCAGCAGCAGCTGTAGAGAGATACAGGCTGCCACCCCTGGGGACGCCCGGGAGGTG 60
QY 389 ggaagcgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 448
Db 61 GGAGAGCCAGGAGGAGCCTGGTGGTCTCTTACTGTGCTCAGAGAGACCGAGCTTCTGTTGT 120
QY 449 gaaagcgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 508
Db 121 GAAAGCAGGAGCTTCCCTCCGCGCGCTCCGAGCTTTTGGCTTTGGGAAAGACCG 180
QY 509 gacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 568
Db 181 GACGAGCTGTGTGTGTGATCTGTTGTTTCTGTTCTGTGTCTTGTCTTGTGCGCT 240
QY 569 cctgtctacagtttaataatgagacagacagcagcagcagcagcagcagcagcagcagcagc 628
Db 241 CCTGTCTACAGTTTAAATATGAGACAGACAGTACAGACCCCTTAATGTTAACTCTCA 300
QY 629 ccattgactgaagtttagatccagagcagcagcagcagcagcagcagcagcagcagcagc 688
Db 301 CCATTGACTGAAGTTAGATTCAGGCGCTCATATTTGTGAGTTACAGTTAAGAGGAC 360
QY 689 ttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 748
Db 361 TTGGCAGACTTCTGTGCTCTGCTCAATGAGCCACATTCGATGTTGATGAGCCATCAGAGG 420
QY 749 gaccttaactcgaatatactcctgctgttaagcgaatatttcaacagcagcagcagcagc 808
Db 421 GACCTTAAATTCGAATTAATCTGCTGCTGTTAAGGCAATCAATTTCAACTGAGACCCAG 480
QY 809 ctctcaatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 868
Db 481 CTCTCAATCAGCAGCAGGAGCCATATCTTACGTGCAAGACTTGCGAAGATCTCTC 540
QY 869 gcaatggttaacacatgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 928
Db 541 GCCATGGGTTAAACCATGTGCTTAATTAACCAAGAAAGCCAGGTCGCCGATCTGCTCT 600
QY 929 ttgagcgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 988
Db 601 TGGGGAAGAAACAAACACTGCGCGCAAAAGATGACGCTCTC-TCTCTGTAATCTACCCG 659
QY 989 agatcgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1048
Db 660 AGATCGAGAGGAGCGCGGATTTGGCGGAGAACCCCAACCTTCTCCACCCCTTATTCAG 719
QY 1049 cacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1107
Db 720 CACAGGCTCTGTGAGGGAGCCCTCTGCCCTCTGAGCTCGGCTGGTGGAGGAGCTG 779
QY 1108 ctgacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1167
Db 780 CTGCGCGAGCTGAGAGCCGAGAGGCGCACCCCGAGCGAGACACAGATCGCGAAT 839
QY 1168 tacgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1227
Db 840 TACCGGTGGGACCATATGCGCTCCCATGCGGGGGGCGCAATGACAGCCCTCCAGATTT 899
QY 1228 ggcctctctctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1287
Db 900 GGCCTCTTCTCTCAGATCTATATATGGAATAACATCAATCCCTTTCTCGGAGG 959
QY 1288 atcccaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1347
Db 960 ATCCCAACAGCCTCACGGGGTGGTGGAGTCTTATGTTCTTACACAGCCTACTTGGG 1019
QY 1348 atgatttcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1407
Db 1020 ATGATTTCAGCAGCTGCTGACAGACTCTTTCACAGCCGAGAGGAGGAGGAAATTCGT 1079
QY 1408 tagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1467

|||||
Db 5459 TTATCCCTTACTGGCTGATTTATGACCTGATACGGGCTGCACTGTAATATACACCTCGA 5518
QY 5845 ggtgttcctccatagaagcaacctggtgctgaactgacattctgctccgactgathaac 5904
Db 5519 GGTGTCTCCTCTAGAGCAGCTGCTGCTGAACCTGATTTCTCCCTCATTTGATTAAC 5578
QY 5905 cccgtgttaaaagcacacctcccaactagtcgtatgtgtgtctgtctgtccca 5964
Db 5579 CCCGCTTTAAAGCACACCTCCCAACCTAGTCGTAAGTATGAGGTTATTCCTCGCCA 5638
QY 5965 ggcacagaaagaaataactgtggtggtctctggaactccctctgtagagatgagc 6024
Db 5639 GGCACAGAAAGAAAGAAATACTGTGGGGTTCTAGAGAACTCTTCTTAGAGATGAGAC 5698
QY 6025 tgcgtcacctcccaagatggaactggaatgycgaactctctccagaacggtgtaaa 6084
Db 5699 TGCCTCACTCCACGATGAGACTGGAATGGCCGATCTCTCCAGGACCGGTTAAA 5758
QY 6085 ttctccttgcattccatccgscgggaagtacaaatgataaactataaagataag 6144
Db 5759 TTCTCCTTTGCAATTCGGCCGGGCAAGTACAAAGTGAATATTAAGATAG 5818
QY 6145 agctgtcccatcagaactagattatcctaaagataagttcactg--aaaggaacag 6202
Db 5819 AGCTGCTCCCATCAGACTTATGATTAAGATTAAGTTCACTGAAAAAGAAACAG 5878
QY 6203 gaaatattcaaaagtgataataatgataagctggtggaataagttttat-tatatgc 6261
Db 5879 GAAATATATTAAGATGATTAAGTATGATGAGTGGGAAATGTTTTTATTAATATGAGC 5938
QY 6262 ggggaagcaggtgctcaacttaacatccgcttagagataagacgggaacacccct 6321
Db 5939 GGGGAGCAGGCTCCTTAACTTACCATTCGCTTAGGATAGAGACGGGACAGAACCCCT 5998
QY 6322 gtggcaatggaacccgataaagtaactggtgaacagggccccggccctggaacacg 6381
Db 5999 GTGGCAGTGGGACCCGATTAAGTACTGCTGAACAGGGGCCCCGGCCCTGAGACACCG 6058
QY 6382 cataactgcccgtgcccacataaactcgtcgtcgcgcctgacataaacaagccgctagc 6441
Db 6059 CATAACTTGGCGGTGCCCAATTAACCTGCTGGCGCTGACATTAACACACCCCTAGAC 6118
QY 6442 aacgtaccatgattgattcctcaacaacgcctagaactcccaagtggtccgt 6501
Db 6119 AACGCTACCTGGATTGATTCCTACACAGCCTTAAACCTCCAGGTTCTCTGTT 6178
QY 6502 aagaagacagagactctcactcactcaacagagagcttccaagcactcaactcaac 6561
Db 6179 AAGACAGGACAGAGCTTTCAGTCTCATCCAGGAGGCTTCCAAAGCCATCACTCCACC 6238
QY 6562 gaacctgacgaactctctgtgtgtcgttgcctacacacagagccctcttataag 6621
Db 6239 GACCTGATGCCACTTCTTCTTGTGGCTTGTCTATCTCCAGGCGCTCTTATTATAG 6298
QY 6622 gggatggttaagaaagaaatlaaagtgaacaaagacatagaatcaatgataag 6681
Db 6299 GGGATGGCTTAAGAGAGAAATTCATGTGACCAAGAGCATTAAGAAATGTCATGG 6358
QY 6682 ggggtccgaataaagctcactcaactgaagttccgggaaaggggaactgataagaaa 6741
Db 6359 GGGTCCGAAATTAAGCTTACCTCCTCACTGAGTTTCCGGAGGGGACATGATGAGAAA 6418
QY 6742 gctcccccatacccaacaacacttgcctatagctggtttatagcagagcctgaa 6801
Db 6419 GCTCCCATCCACCAACACCTTGTGCTATAGTATGCTGTTTATGACAGAGGCTCAAGA 6478
QY 6802 aatcagattatagctcgtgtatataaaggtggtggaatgacaactaggtgttaacccc 6861
Db 6479 AATCGATTTATGATCTGTTTATACAGGTGTTGAGCATACTGGGTTAAACCCC 6558
QY 6862 tgtgttccacctcagttcaacacatccaaagattctgtgtcaatggtccaaatcgtc 6921
|||||

Db 6539 TGTGTTCTCCTCCAGTCTTCAACCAATCCAAAGATTTCTGTGTCATGGTCCAAATCCTC 6598
QY 6922 ccccgagttactacacatccatgaagaagtgtcctttagatgaatgactatcgtataac 6981
Db 6599 CCCGAGTGTACTACCATCTGAGAGAGTGGTCTTGATGAATATGACATGAGTATTAAC 6658
QY 6982 cgacaaagaaagaaacccglatcccttaccctagctgtlaatgctcgtatagaagcggc 7041
Db 6659 CGACAAAAGAGAAACCCGATATCCCTTACCTAGCTAGCTAATGCTCGGATTAGGACGGCC 6718
QY 7042 gtlgycgtagaacagggacagctgcctcgtatcaacaggaacacagcagctagaagaa 7101
Db 6719 GTTGGCGTAGAAACAGAGGACAGCTGCCCTGATCAGACAGCACAGAGCTTAGAGAAAGA 6778
QY 7102 ctgtgtgaactacatgycgcacatgaagaagttccggagccttaagagagctgaac 7161
Db 6779 CTTGGTACCTAATCGGCTGATGACAGAGATCTCCGAGCTTAAAGAGTCTGTTAC 6838
QY 7162 aacctagaagatccctcactctcttctgtagagtggtctcaagaacccggagggatga 7221
Db 6839 AACCTAGAGAGAGTCCCTGACTTCTTGTGTAAGTGTCTTAAGAACCGAGGGGATTA 6898
QY 7222 gactgctgttctcaagagaaggtggtatgtcagccttaagaagaatggtgttc 7281
Db 6899 GATCTCTCTTCTTAAGAGAGAGTGGTATGTGACGCTTAAACAAAGATGCTTTC 6958
QY 7282 tatgtatgaactcagagacatcagagactcactaacaacttgaagaaatgataag 7341
Db 6959 TATGTATGATCCTCAGAGAGCATCAGAGACTCCATAGCAAGCTTAGAAGAAAGTTAGAG 7018
QY 7342 aggcgtcgaagggaaagagagctgacacaggggtgttgaagatglttcaacaggtc 7401
Db 7019 AGGCGTCGAAGGGAAAGAGAGGCTGACAGGGGTGTTGAAGAGATGTTCAACAGTCT 7078
QY 7402 ccttgatgacacacctgtcttctgtctgacggggcccttagtctgtctccgtga 7461
Db 7079 CTTTGGATGACCACTCTCTCTGCTGACGGGACCCCTAATCTCTGCTCTGTTA 7138
QY 7462 ctacagttggcctgcttaataatgaagttgtgtgctgtgttagaagaacagtgat 7521
Db 7139 CTTACAGTTGGCTTGTCTTAATTAATAGTTGTTGCTTCTTGTAGACAGAGGAT 7198
QY 7522 gcaactcagatcactgacttaggcaacagttacaaagcctctgacaaagagaact 7581
Db 7199 GCAGTCCAGATCAMPAGTACTTAGCAACAGTACCAAGGCTTGTGACCAAGAGAACT 7258
QY 7582 gaacctagcctcccaagttctagaatgaactatcaacaagaagaagtggtgaag 7641
Db 7259 GACCTTAGCCTTCCAGATTAAGATTAACTATTAAACAAAGAAAGTGGGAAAG 7318
QY 7642 aaagatgaataatgcaacttaacccctccagaacccagaagttataaagaactaa 7701
Db 7319 AAAGGATGAATAATGCAACTTAACCTTCCGAACCCAGAGAAAGTAAATAAAACTTAA 7378
QY 7702 tgcgcccgaaatlaacaacctgtgctgctgcaagtaaatgtagaagttcaactccta 7761
Db 7379 TGCCCCGAATTCACAGACCTGCTGCTGCAGTAAATAGATGAGAGGTACACTCTTA 7438
QY 7762 ttgttccagggcctgtctatccgtcgtcgaatgaataaagaagaatgagttgactaag 7821
Db 7439 TTGTTCCAGGCGCTGCTATCTGCTTAATTAAGATTAACGGAATGAGCTGATATG 7498
QY 7822 ctatctggaattctgtaaaactgactgacacatagaagaattgattacacattgacagc 7881
Db 7499 CTTATCTGGATTCTGTAATCACTGACGACATTAAGAGATGATTACCATTTGACAG 7558
QY 7882 cctagtgacctatcactgcaactgtaactgtcactgcccagaagccacagcagatgcga 7941
Db 7559 CTTAGTACCTATCTCAATGCAATCTGTCACCTGCCAGGACCCAGCATGATCGGA 7618
QY 7942 cctccggagctattttaaaatgattgltccacaggaagcggggtcctgattttaaat 8001
Db 7619 CTTCCGAGGCTATTTAATTAATGATGTGCTCACGAGGCGGCTCTGATATTATTAAT 7678
|||||

[illegible]

QY	1380	acaccgagagcgagagagaaatttcgtttaagagctcagaanaaatggtcccgggcgac	1439
Db	1907	ACAACCGAGGCGCAGAGACAAATTCCTTTAGAGGCTAGAAAAAATGTTCTCGGGCCGAC	1966
QY	1440	ggcgacaccacgcagctgtgcaaaaatlgagatlgaaatlygaaattcccttgactcgcccggt	1499
Db	1967	GGGCGACCCACGACGAGTTGCCAAATAGAGATTGACATGGATTCCCTTGACGCCCCGT	2026
QY	1500	ttggactacaacacgcgtctgaagtgaaagagagcttgaaaaattcalcgccaggtcttggtg	1559
Db	2027	TGGGAGCTACAACACAGCGCTGGAAGTAGGAGAACGTTGAATATATACGCGCAGGCTCTGGTG	2086
QY	1560	gcgagttcccgaggcgctctaagaacgcgaactaatlttgctcaaggtgaagaagtgatg	1619
Db	2087	GGGGGTCTCGGGGCGCCTCAAGACGGCCCACTAAATTTGGCTAAGGTAGAGAGGTGATG	2146
QY	1620	caggagaccgaacgcagactccctcgtlatlctcttgaaggtcatalgaaagcttcagcgcg	1679
Db	2147	CAGGAGCCGACACGAACCTCCCTGGGTATTTCTTGAGAGGCTCATGGAAGCCTTCAGCGGG	2206
QY	1680	ttcacccctlttgatctcctaccatcaagagcccaaaaagctcagtggtcgccgtgctcaat	1739
Db	2207	TTACACCCCTTTTGATTCCTACTACTCAGAGGCCCAAAAGCCTCAGTGGCCTTGCCCTTCAT	2266
QY	1740	ggcgagttcggtcgtctcgtgatalcaggaagaacttcagagcttgaagaggtttacagaagct	1799
Db	2267	GGGCGTCTGGCTCTGGATATCACAAAAAATTCACAGACTGGAAGGGTTACAGGAGCT	2326
QY	1800	gagttacgttgctctagtgagagagcgcaagaagtgctatcacagaagagagacagaag	1859
Db	2327	GAGTTACCTGATCTTAGTGAAGAGGCGCAGAGAGGTATTACAGAGGGAGACAGAGAG	2386
QY	1860	gagaaagacagagaaaaagaagaagagagaaagaagaagagaaagcgtgatagacgtg	1919
Db	2387	GAGAAAGCAACAGAAAAAAGAAAGGAGAGAAAGAGAGAAAGGAGAAAGACGTGATGACGG	2446
QY	1920	caagaagaactlttgactaagatcttggtccgagtggttgaaaggaagacagcagggag	1979
Db	2447	CAAGCAAGAATTTTGACTTAAGATCTTGCGCGAGTGTTGAAGGGAAGAGCGACGGAG	2506
QY	1980	agaagagagagttttaagaaaatttaagtgagccctacagacgttcaggaagaccttggaat	2039
Db	2507	AGAGGAGACAGATTTTATAGAAAATTAATGATGCGCCCTAGACAGTCAAGGAACTGGCAAT	2566
QY	2040	aggaccaccaatcgcgaacagagccagtggtgctatctgaagaanaaagacacatlygacaag	2099
Db	2567	AGGAGCCCACTCGACAAAGACAGACAGTGTGCTATTGTAAAGAAAAAGACACTGGGCAAG	2626
QY	2100	aactgtcccaagaagggaaaaacaagaacgaagctctcagctctcgaagaagataaagt	2159
Db	2627	AACCTGCCCAAGAAAGGGAACCAAAAGCCAAAGAGTCTGTAGCTGTGAAGAGATTAAGAT	2686
QY	2160	tagggagaaacgggtltcgagcccccctcccgagcccaaggttaactttgaaggtgaaggg	2219
Db	2687	TAGGGAGACCGGGTTGGACCCCTCCCGACCCAGGATTAACCTTGAAAGGTGAAGGGG	2746
QY	2220	caacacagttgagttcccttgattgataccggaagcgagacatcagttcgtgtaaacacata	2279
Db	2747	CACACAGTTGAGTTCCTGCTTGATACCGGAGCGGACATTCAGTGTGCTACAAACATTA	2806
QY	2280	ggaaaaactaaaaaanaaaatcctgtgtgataggtgtgcacagcgcaagcgaglatcca	2339
Db	2807	GGAAGACTTAAAGAAAAAATCCTGGGTATGCGTGCCACAGGGCAAGCGCATATCA	2866
QY	2340	tgagactaccgaagaacggttgacttgagagtgaggacgggtaacccactcgtttctgtgc	2399
Db	2867	TGGACTTACCCAAACACGTTGATCTTGAGGTGGAGGAGGATTAACCCACTCGTTTGGTC	2926
QY	2400	atcccttgatggccagtaaccctctagtgatagagacttaactacaagaatggagagctcaa	2459
Db	2927	ATCCCTGAGTGCACAGTACCCCTTCTTAGTGAAGAGATTACTGACCAAGATGGGAGCTCAA	2986
QY	2460	attcttltgacaagaagacagaaigtgtgataatacaaaaccaatcgtgtgtg	2519

gene	CDS	/translation="MGCTVTPPLSTIDHWTEYKSAHNLSVOVKRQPCWOTFCVSEMP TFDVGWSEGTENSELLAVKAILIFQGPSHPDQETITLIMODLADDEPPWPKMLN KPRKPGRIALAEKNSHSAEKVPSHPIELEEPPAWEPDPSVPPYPAGAGK PSAPGAPAVEGPAAGTARRGATPERTDEIATLPTETGPIPGQLOPLQSS ADLYNMKTHPPESEDPQRLGLVSELMFESHQPTDCOOLLOTLTETTERKILIEA RKVNGADGRPTOLONEIDMGPLTRGCMYNTAEGRESLITRYOLVALRGSARPA TNLAKEVWOGPREPPSVLEERMEFRPPTDPSFQAKSMALAIIGOSALDIR KLIORLEGLADAEALHDLVKEAEKVTIKRETEEBERQKEREERERERRRKRREKMTIR KILAAVEGKSNRERERDFRRKINSGRPOGSLNRRPLDKOCAYCEKKGHWARDCK KNGKGLVALLEEDKD"	
		2728. .6309 /gene="pol" <2728. .6309 /note="complete pol ORF, gag-pol is expressed as a polyprotein" /codon_start=1 /product="polymerase" /protein_id="CAC39617.1" /db_xref="GI:14275839"	
gene	CDS	/translation="GRGSDPLPEPRPTLVESQPVFVLDTGAKHSVLLQPIKID KRSWMGATGQOQYPTWTRRTVDLGVAVHSFLVTECPALIGRLILKMGQISE BOGRPEVSANKRITVLLDLDEYRLVSPVKPDONIQWLEDFPQAMETAGMGLA KQPPQVQLKASAPVSQYPLSKARERIRHVORLIDQGLIVVQSPWMTPLP VRKPGTNDYRPVODLREYNKRKVDIHPTVPVYMLCALPORSMTYVLDLKAFFEL RLHPTSOPLAFEMWRDQAGRTGOLTWRLDQCKNSPTIJDDEALHRLDNFRIOHQ VTLLQYVDDLACATKQDCELEGTALLLELSDGQASAKKMOICREYVIGSIR GGORWLEARRRIVQIPAPTTAKOVEFEJGAFCKMLGEATLAPLPLTKEG EFSWAPBEHOKAFDAIKALLSAPALAPDVITLPEDEKGAVALVUTLWPKR PVAYSKLIDVASGMPICLKAIAVALIVADAKTLGOMITVIAHALENTIROP DRMTNARMTYOSLLTERVTFAPALNATILPETDEPVDHGOHLEIETGVA KDLIDPLTGMILMPTDSSSYWEGKAGAAVADGRTIMASSEDSQAEELM ALTOALRAEKSINTYDSRYATATVHAITKOGILTSAREIKKEILSTLE ALHDKRLAIIHCEHOKAKOPISRGONADPRAVKAQAOANILPMIETKRAPSGO YLDEMOEIKRIDFSETEGCTYSDKELTPHEGLEEVQOIHRLHLSKTLQOL VATSEYHVLRFVADVSVKHCPOVLNANPSTRIPKRLRGSHPAHMEVDETEK PAKYGNKYLIVDFVDSGWEAVYPKKSTYVAKLILEETFPFRGPIKTVSNGPA FVAOVSQGLAKILGIDMKLHCAYRPOSGOVERMNRKTETLTLTETGINDMALL PVLFRVNTPGQGLTPEYELLGSPPLAELAFASHDVLLSOPLSRKLALMVRQ RAMKQLEAYSQGDQVPHRFQVGSYVRRHRGNLETPRKSGYIVLITTPAVKE GTPVHSHKAPPPDSGMAKTEPNLKLRLHRYVPYVNNSS"	
		6185. .8149 /gene="env" 6185. .8149 /gene="env" /codon_start=1 /product="envelope" /protein_id="CAC39618.1" /db_xref="GI:14275840"	
gene	CDS	/translation="MHPILSRHRLPIRGKPKRLKIPLSFASIAMFLTLTPQVNC RLVDSNHSKPLSTLWLTDSGTGININSTGEALGTMPPELYVCLRSVJGLNDQA PPDYLRAGFYVCPGPNNEVEYCGNPGKSCITSDNKMKNWPSOODRYS FYNNPTSTNOFTYGRKMKDQORQOKYRNKOISCHSLDLYDKLSTPEGKQDNTS KWNNGISWGLVYGGSGRKGSLVITRLIETOMPEVAIGPNKLAOGPIIDORP SNPDSYNTSGSVTEPNITITRTAKLSLIGAFQALNSTPATSSCWLCACTEP PYEGMARGKENVTRKEHDDCTGWSOKLILEVSGGTCIGMVPBHQHCLNTEA FNRTSEGYLVGYDRMAACNTGTPCAGTVFNQTKFCVAVQVLPVYVYKEAVL DEVDYRNKRPKEPISITLAVMLGLGVAVGTGAALITGQOLEJKNHRLVTE DIOALEKSYSNLEESTLSSEYVIONRGLDILLFKEGGICVAKLECCCFVDHSAI RDSMSRLRELRKREKEFAGQGFEGFNPNSPWTTLTSLALGTPLVILLLLTYGPC LINRFAVFEVYSAVRIMVLRQYGLPS"	
		8216. .8918 8216. .8918 /LTR /polya_signal /BSP_COUNT 2435 a 2174 c 2253 g 2056 t /ORIGIN	
Query Match	Best Local Similarity 87.9%; Pred. No. 0; Matches 7332; Conservative 0; Mismatches 798; Indels 210; Gaps 15;	79.1%; Score 6493.2; DB 14; Length 8918; 1 gtggtgtaagcgtgtggcccgagcgctgtgataaaatccctctgtgttcgctc 60 	
		562 GTGAGCGNAGCAGTGTGAGCCGCCAGCGGCTCGGATTAATAATCTTGTCTGTTGCATC 621 QY 61 aagacgcctctcgtgagtgatttgggggtgcgcctctccgagccgagcagggagat 120 Db 622 AAGACCCCTTCGTGAGTATTTGGGGTGTCCCTCTTCGAGTGAAGCAGAGAGGAT 681 QY 121 tgtctcttaacgtccttcaatttgggtgcgttggcgggaaatccttgcagaccctta 180 Db 682 TTTAACTCGACGTGGCCTTTCAGTTTGGTGTGCGGGAACCCGCGCATCCCTCA 741 QY 181 caccgagaaacgacttgaaggttaagggatcccccttggaaagt----gtgtgtgtc 236 Db 742 CACCCGAGAACCGACTTGGAGGTAAAGGATCCCTTTGAACTGTGATGTGTGTGTC 801 QY 237 ggcgcgctctcgttccttgaagtgctgttctcgttgaagcgccttcttgcagctg 296 Db 802 GGGCGGCGTCTCTTCTGTGAGTGTCTGTCTTTCGGTGTATGGCCCTTTCGGTTGCA 861 QY 297 tccctcagaacgtaagacttgaagacttgaagcgttgatcaagcagcgttctgaagatacaag 356 Db 862 TCCTCTCAGACCGTAAGGACCTGGGAGCTGTGATCAGCAGACGTGAGGAGATCAG 921 QY 357 gctgcacaccttggggagcgcgcgcggaggttggggagagcgaaggaacgcttgggtctcc 416 Db 922 GCTCCACCCCTGGGGAGCGCCCGGGAGGTGGGAGGACCAAGGACCTGTGTCTTC 981 QY 417 tactcgcgtcagaagacgagcttctgttlttgaagcgaagcttccccctccgagcg 476 Db 982 TTCTGTGCGTCAGAGACCGAGGCTCTGTGTGAAGGAAAGCTTCCCGCCGCGCG 1041 QY 477 tccgactcttcttgccttcttgaagacgagcagcgttgcgttcttgaatcttgc 536 Db 1042 TCCGACCTTTTGGCTGCTGTGTGGAAGACGGGAGCGGCTGCGGTGTGTGATCTTTG 1101 QY 537 gttctcgtcgttcttgccttcttgccttcttgccttcttgccttcttgccttcttgc 596 Db 1102 GTTCTGT 1161 QY 597 aacgtgactacccccccttgaatttgaacttcgacacatcgaagactgaagtaagtaagcagcct 656 Db 1162 ACGGAGACGACCCCTCTTGAATTCACCTCAGCATTCGAGTAAGTAATCCAGGCGCT 1221 QY 657 cataatlttgatcagtttaagaagagagacttgcgaagacttcttgcctcgaatg 716 Db 1222 CATATTTGTGACGTTCAAGTAAAGGAGGACCTTGGCAGACTTCTGTGTGTAATGG 1281 QY 717 ccaacatcgaatgttgaatgacatcaagaagagacutlaatttcaaatatcctgct 776 Db 1282 CGGACATTCGATGTTGATGGCCATCAGAGGGACCTTATTTCTGAGATTAATCCTGGCT 1341 QY 777 gtaaggaatcaatlttcaagacttgaagcgcgctcctcctgatacgaagacccatatac 836 Db 1342 GTTAAGCAATTAATTTTACAGACTGACCCGCTCTCATCCCGATCAGAGACCCCTATATC 1401 QY 837 cttacgttgcgaagatttgcagaagaatactctcgcataaggtttaaaccatlgctaaataa 896 Db 1402 CTTACGTGCGCAAGATTGTGCAAGAGATCCTCGGCATGAGGTTAACTTTGGCTGAATAG 1461 QY 897 ccaagaaacgagtcgcccgaaatccttgccttcttgaagaaacaaacacactcggcgaa 956 Db 1462 CCAAGAAACCGAGTCCCGCAATTCGCTCTTGAAGAAACAAACACATCGGCTGAA 1521 QY 957 aagtcgagccctcttctcgtatcaaccgagatcgaagagccgcgacttggcgga 1016 Db 1522 AAGTCAAGCCCTC-TCCATATATCTACCCAGATTAAGGAGCGCGGCTTGGCCGA 1580 QY 1017 acccaacccttctcccccaccccttaccagacagaaggtgctgtgaagaga-cctctgc 1075 Db 1581 ACCCAATCTGTCCCAACCCCTTATCCGACAAAGGCTGTGAGAGGACCTCTGC 1640 QY 1076 ccctcctgagtccttgcgttgaagagacttgcgcgagactcggagacggagagcgc 1135 Db 1641 CCTCTCTGAGGCTCCGGCGGTGAGAGGACCTCTCAGAGGACTCGAGACCGGAGGCGCG 1700	

QY 1136 caccccgagcggacagacgagatcgcgataltaaccgtcgcacactatggcccccat 1195
|||||
Db 1701 CACCCGGAGGCGGACAGACGAGATGCGACATTTACCGGTGGGACAGTACGGCCCTCCCAT 1760
QY 1136 gccagggggccaatttgagccccctcagatattggcccttctctctcagatctctataa 1255
|||||
Db 1761 ACCGGGGGGCCAAATTGACGCCCTCCACGATTTGGCCCTTTCTCTGACAGTCTCTATAA 1820
QY 1256 ttggaataactaacatcccccttctcggaggatcccccaacgctcaggggttggtgga 1315
|||||
Db 1821 TTGGAAACTAACCATCCCTCTTCTCGAGAGATCCCAACGCCCTCACGGGGTTGGTGA 1880
QY 1316 gtccctatgtctctcaccagcctacttggatgatgttcaaacagctctgcacagacat 1375
|||||
Db 1881 GTCCCTTATGTTCTCTACACGCTACTTGGATGATGTCACACAGCTGCTGCGACGACAT 1940
QY 1376 cttaacaaacgagagagcagagaatctctgtttagaggtctagaaaaatgttccctgggc 1435
|||||
Db 1941 CTTCAACAACGAGAGCGCAAGAAATTTGTTAGAGCTAGAAAAAATGTTCTCTGGGCGC 2000
QY 1436 cgaacggcgacccacgcagctgtcaaatgagattgacatggatttcccttgactggcc 1495
|||||
Db 2001 CGACGGGGGACCCACGCGATTGCAAAATGAGATTGACATGGGATTTCCCTTGACTGCGCC 2060
QY 1496 cgglttggaactaacaacgcgctgaagtgaggagagcttgaataatctatcgcagcctc 1555
|||||
Db 2061 CGGTTGGGACTACAAACGCGCTGAAAGTAGGAGACCTTGAATAATCTATCGCCAGCGTCT 2120
QY 1556 ggtggcgggtctcggggcgccctcaagaacgcccacactaatttggcclaagtgagaagt 1615
|||||
Db 2121 GGTGGCGGGTCTCCGGGGCGCTCTAAGACGGCCCACTACCTTGGCTTAAGGTAGAGAGGT 2180
QY 1616 gatcagggagcagacgaacactccctcgatattcttgaaggtctatgaaagccttcag 1675
|||||
Db 2181 GATCAGAGGACCGAATGACCTCCCTCAGTTTCTTGAAGGCTCATGSAAGCCTTTCAG 2240
QY 1676 gcggttaaccccttctgtactctactcaagaagcccgaagaaagcctcagttggccctggcct 1735
|||||
Db 2241 GCGGTTACCCCTTTTATCTCTACCTCGAGGCTCAAGAAAGCCCAATGGCTCTGCGCTT 2200
QY 1736 catgggcagctcggcctctgtatctcaggaagaactctcagagcttgaaggtctaaaga 1795
|||||
Db 2301 CATAGGACATCAACCCCTGATATTCAGAAAGAACTTTCAGAGACTGGAAGGCTTACAGGA 2360
QY 1796 ggcgtgagttacgtgatctagtgagagagcagagaaggtgtatctacagaagggagacaga 1855
|||||
Db 2361 GGCTGAGTTACATGATCTAGTGAAGGAGCGAGAGAAAGTCTATTACAAAAGGAGACACA 2420
QY 1856 agagggagaagacagagaagaagaagagagagagaagaagggaggaagacgtgataag 1915
|||||
Db 2421 AGAAGAAAGGAGCAAAAGAAAGAGAGAAAGAGAGAAAGGAGGAAAGGAGCTAAATA 2480
QY 1916 acggcagaagaagaatttgaactaaga tcttggccgcaaggtgttgaagggagaagagcgag 1975
|||||
Db 2481 ACGGCAAGAGAAATTTGACTAAGATCTTGGCTGCAAGTGCTTGAAGGAAAGCAATAG 2540
QY 1976 gggagagagagagatlttagaanaa ttagtcaagggcccttagacagctcagggaaactggg 2035
|||||
Db 2541 GGAAGAAAGAGAGATTTTAGMAATTTAGTCAGGCGCTTAGACAGTCAAGGAAACCTGGG 2600
QY 2036 caatagagaccccaactcgaaagaagcagtgctgtatgttgaagaaaaagagcaactgggc 2095
|||||
Db 2601 CAATAGAGACCCCACTCGAAGAGACCAATGTGCAATTTTAAAGAAAAAGGACACTGGGC 2660
QY 2096 aaggaacttgcaccaagaagaagaacaagaagacgaaggtcagctcctagtaagaagataa 2155
|||||
Db 2661 AAGGAGACTCCCAAGAAAGGAAACAAAGAGACTGAAGGCTTAAAGCTCTGGAAGAAATTA 2720
QY 2156 agattagggagagcgggttctggaccctcccgagcccaagggttaactttgaaggttga 2215
|||||
Db 2721 AAGACTAGGAGACGGGTTTCGACCCCTCCCGAGGCCAAGGTAACCTTAAAGGTGA 2780

QY 2216 ggggcaaacagttgagttctcgttgataccggaagcgagcattcaagtctctataacc 2275
|||||
Db 2781 GGGCAACCAAGTTGAGTTCTGTTGATACGGAGGAAACATTTCACTACTACAGCC 2840
QY 2276 attagaaacataaagaanaaaatcccttgggtgagtggtgtccacagggcaagcagta 2335
|||||
Db 2841 ATTAGAAACCTAAAGATTAATAATCTCGGTGATGGGTGCGCACAGGGCAACACATTA 2900
QY 2336 tcaatgtactacccgaagaacgcttgacttggaggtgggaacgggttaaccactcgttct 2395
|||||
Db 2901 TCCATGGACATACCCGGAAGAACATGTAATGGAGTGGGAGCGGATTAACCATCGTTCT 2960
QY 2396 ggtcaccctcagtgagtcgcagtaacccttctagtagaagcttacttaccagaatggagc 2455
|||||
Db 2961 GGTATACCTGAGTGGCCAGACACCTCTTAAAGTAAAGACTTACTTACCAATGGGAGC 3020
QY 2456 tcaattctcttgaacaaagaagaacgaagtgcttggatgaataacaaacccatcag 2515
|||||
Db 3021 ACAATTTCTTTGAACCAAGGGAACCAAGAGTGTGCAATTAACCAATCTACTCT 3080
QY 2516 gttgaacccccaatttagatgatacgaactatctcccaagtaaaccttgatca 2575
|||||
Db 3081 GTTGACCCCTCCAAATTAAGATGACGAATATCGACTATATCTCCCTAGTAAACCTGATCA 3140
QY 2576 agatatacagctcgttggagcagtttcccaagccttgggagaagaacgagagatgg 2635
|||||
Db 3141 AATATACAAATTTCTGTTGGAACAGTTTCCCAACCTTGAGGCAACCGGAGATGGG 3200
QY 2636 lttggcaagaagaagttcccccacaggttattcaactgaagggcagtgtaacacagatc 2695
|||||
Db 3201 TTTGGCAAGCAAGTTCCCCACAGGATTAATCACTGAAGGCGCACTGCTGACCAAGTTC 3260
QY 2696 agtcagacagtagtcccccttggatgagagagcttcggaagaatttggccgcatgtcaag 2755
|||||
Db 3261 AGTCAGACAGTACCCCTTGAGTAAAGAGCTGGAAGAAATTCGCGCATGTTCAAG 3320
QY 2756 attaatcaacagggcatccctagctcgttcccaactcccccttggaaatctccctgtacc 2815
|||||
Db 3321 ATTAAATCAACAGGGGCAATCTAGTTCTGTTCAATCTCCCTGGAATCTCCCTGCTAGC 3380
QY 2816 ggttgaagaagccttggacccaatgattatcgacaggtacagagacttgaagaagttcaata 2875
|||||
Db 3381 GGTTAAGAAAGCTCGGAGCAATTAATGACTATCGACAGACTGAGACTGAGAGAGTCAATA 3440
QY 2876 aaggttgcagacataacccaaggttccgaacccctataactctttagcgcctccc 2935
|||||
Db 3441 ACGGTTGAGATGATACACCAACAGTCCGAAACCTTATTAACCTTGTGCTCC 3500
QY 2936 gcttgaagcagacttgttacaagatattgacttaaaagatcccttctctcctgagatt 2995
|||||
Db 3501 ACCCCACGAGAGCTGTATACGATTTGAGACTTAAAGATCCCTTCTTCTGCTGAGACT 3560
QY 2996 aacacccactagccaacacttlttgccttcgaatgagagatcccaagttacaggaagaac 3055
|||||
Db 3561 ACATCCCACTGACCAACCACTTTTTCCTTGAATGAGAGATCCAGGTGGGGAAGAAC 3620
QY 3056 cgggcagctcaacttggacccgacacttgcgaaggggttcaagaagctcccccacactttga 3115
|||||
Db 3621 CGGGAGGTCACCTTGGATGACACTGCCCAAGGGTTCAAAAACCTCCCGACCACTTTTGA 3680
QY 3116 cgaagccctacacagggaccttggccaacttcaagatcccaacacccctcaagttgacctcct 3175
|||||
Db 3681 CGAAGCCCTTACAGAGAGCTGGCCACTTCAAGATCCCAACACCCGAGGTGAGACCTCTCT 3740
QY 3176 ccaatagcttgaatgactgtcttcttggcgaggacccaacaaacaggaactgttgaaggtac 3235
|||||
Db 3741 CCAGTACGTGATGATACCTGCTTGGGGGAGGACCAACCAAGAGACTGCTTAAAGGTAC 3800
QY 3236 gaagcaactatgttgaatgttgcacttaggttaagggctcttgccttaagaagggcca 3295
|||||
Db 3801 GAAGGCACTACTGCTGGAATTTGTGACCTAGGCTACAGAGCTCCCGCTAAGAAAGGCCCA 3860
QY 3296 gattgcagagagaglaacataacttgggttacaagtttgcgggagcgagcagatgct 3355

```
|||||
Db 3861 GATTTCAGAGAGAGAGTAACATCTTGGGCTACCTTTGGGGGGGCGGAGAGATGGCT 3920
Qy 3356 gaaggaagcaggaagaaatctgtagtccagataccgcccacacacgaacaaagt 3415
Db 3921 GAGGAGAGGACGGAAGAACTAGTCCAGATACCGGCCCCMACACAGCCAAACAACT 3980
Qy 3416 gaagaagttttgggagacgtgtagtttggagacgtgtagtccggggttggagcctt 3475
Db 3981 GAGAGAGTCTTTGGGACACCTGGATTTTGGACAGCTGTGATCCGGGGTTTGGACCTT 4040
Qy 3476 agagccccaactaacccgtaacccaagaaagggaattctcctgggtcctctaaga 3535
Db 4041 AGCAGCCCACTCTACCCATTAACCAAGAAAAGGGGAATCTCTCGGGCTCTTAGCA 4100
Qy 3536 ccagaaggaacttgaatgatacaaaaagccctgtcgtgagcgaactgtcttgccctcc 3595
Db 4101 CCAGAAAGCATTTGATGCTATCAAAAAAGCCCTGTGAGGGCACCTGCTTGCCCTCC 4160
Qy 3596 tgacgtactaaaccccttaccctttagtggatgaagcgtaaaggagtagcccgagagt 3655
Db 4161 TGAGGTACCTPAACCCCTTTACCTTTATGTGATGAGCGTAAAGGAGTACCCGGGAGT 4220
Qy 3656 tttaaccnaaacccagacatggaagagagacgtgtgtctactactgtlcaagaagactga 3715
Db 4221 TTTAACCCAAACTGTAGACCATGAGAGAGACCTGTGCTACCTGTCAAGAAAGCTCGA 4280
Qy 3716 tccctgtagcagtagtgtgcccgtatgtctgaaggtatcgcagcgtgtgcaactgtgt 3775
Db 4281 TCCTGTAGCCAGTGTGTGGCCCATATGCTGAAGGGCTATGCGAGCTGTGCGCATCTAGT 4340
Qy 3776 caaggaagcgtgacaatctgacttggagcagaatataactgtaataagcccccatgatt 3835
Db 4341 CAAAGGAGCTGACAAATTTGATTTGGAGACGAATATACGTGTATTAACCCCATGCTT 4400
Qy 3836 ggaagaacatgcttcgagagcccccagacagatgtagtaccacgcgcgcga tgaacca 3895
Db 4401 GGAAGACATGCTTGGGACAGCCCGACGACCGATGATGACCAAGCCCGCATGACCCCTA 4460
Qy 3896 tcaaaagcgtcttctcaagaagaagtgtaacttgcgtccacagccgcttcaaccctgc 3955
Db 4461 TCAAAAGCTGCTTCTCAAGAGAGGGTACGTTGCTCCACACCCCTCTCAACCCCTGC 4520
Qy 3956 caacctctgctgaagaagactga tgaacaggtactacatgattgcatcaactatgat 4015
Db 4521 CACTCTCTGCTGAAGAGACTGATGAACAGTACTCATGATTTGCCATCAACTATTGAT 4580
Qy 4016 tgaagaagactgggtccgcaagagccttacagacataccgctgacttggagaagtgtaac 4075
Db 4581 TGAGAGAGCTGGGGTCCGCAAGAGACCTTACAGACATACCGCTGACTGAGAAATGTTAAC 4640
Qy 4076 ctggttcaactgaacgaagaagctatgtgtgtgaagtaagaagtagtgcgtggcagct 4135
Db 4641 CTGGTTCACTAGCGGAAGACACTATATGTGTGAAGGTAAAGAGTATGCTGGGGGGGGGT 4700
Qy 4136 ggttgaagcggagcccgcaagactcttggtgccaagagccttcggaaagaacttaagcgaaa 4195
Db 4701 GGTGACCGGACCCGACGATCTGGGCGACGACGCTGCGGAAGGAATCTTACGACAAAA 4760
Qy 4196 ggttaagctcatggtccctcaagcaagcttgcggtcgtgcggaaggaatccataaacat 4255
Db 4761 GGCTAGCTCATMGGCCCTCACAGCAAGCTTTCGGGCTGGCCGAAGGAAATCATTAACAT 4820
Qy 4256 ttataaggaacgaagtagtcccttctgactgtcaacgtatcacaggggcatactataaca 4315
Db 4821 TTATACAGACAGCAGATGTGCTTTGGCAGTGCACACCTTACACGGGGCCATCTTAAAGCA 4880
Qy 4316 aagggtgtgtcttaactcaagcaggaaggaataaagaacaaagaagaatctcaagcct 4375
Db 4881 AAGGGGGTTGCTTACCTTAGCAGGAGGGAATTAAGAACAAAGAGAAATTTCTAAGCTT 4940
Qy 4376 attagaagccttacaatttgcanaaaaggctagctattatacagtgctcgtgacatcaga 4435
|||||

Db 4941 ATTAGAACCTTACATTTTGCCAAAAAGGCTAGCATTTATACATGCTGCTGCATCAGAA 5000
Qy 4436 agccaagatctcatatctagagaagaccagatggtgaccgggttgcgaagcagcagc 4495
Db 5001 AGCCAAAGATCCCATATCTCAGAGGAGAACCAAGATGCTACCGGGTGGCAAGGAGCAGC 5060
Qy 4496 ccaggtcttaacctctgctcatalaagaacgcccacaaagccccaagaaccagaagaca 4555
Db 5061 CCAAGGTGTAACTTTCGCTATATGATGAAACACCCCAAGCCCGCAGAACCCGGACGACA 5120
Qy 4556 gtacaccttagaagactgccaagagataaaaaagtagaacagttcttgagactccgga 4615
Db 5121 GTACACCCCTAGAAAGCTGGCAAGAAATAAAAAGATPAACAGTCTCTGAACCTCGGA 5180
Qy 4616 ggggaacctgtatactcaatggaagaaatccgtgcccacaaagaaggttagaata 4675
Db 5181 GGGGACCTCTTACTCTCAGATGGGAGAAATCTGCGCCCAAGAAAGGGTTAGAAATA 5240
Qy 4676 tglcaacagatatacgtlctaacccaactagaaactaaacactgcagcagttgttag 4735
Db 5241 TGTCACACAGATACATCGTCTAACCCACCTAGAACTAAACACTGACAGCATGTTGTCAG 5300
Qy 4736 aacatcccttatcatgttctgaagctaccagagatgtgtgtgtgtgtgtgtgtgtgt 4795
Db 5301 AACATCCCTTATCATGTTCTGAGGCTACAGAGAGTGTGACTGTGCTGTGCTGTGCTGTG 5360
Qy 4796 tgtgacctgcagcgtgtgtatgtatctatccctccagaatacctccaggaagaagataag 4855
Db 5361 TGTGCTCTCCAGCTGTGTATGTCTATGCTCTCCAGAAATACCTCCAGAAAGAGCTAAG 5420
Qy 4856 ggaagaccacccaagcgtcactgtggaagtgtgacttcaactgaggtlaaagcggctaaata 4915
Db 5421 GGGAAACCCACGAGCGGCTCACCTGGGAAGTGAAGTCACTGAGGTAAACCCCGCTAAATA 5480
Qy 4916 cggaaacaatatctatgtttttagaaccttttagaagtagtagaagtagtagtagtag 4975
Db 5481 CGGAACCAATATCTATTTGTTTGTGTAGACACTTTTGTAGAGTGGTGAAGGCTTAATTC 5540
Qy 4976 tactaagaagagacttcaacgctgtgtgclaaagaataactggaagaaattttccaag 5035
Db 5541 TACTAAGAAAGACCTTCAACCGTGTGTGCTAAGAAATATCTGAGGAATTTTCCAG 5600
Qy 5036 atttgaatacctaaggaataggttcagacaatgtgtccagcttccgttgcgaagtag 5095
Db 5601 ATTTGAATATCACTAAGGTATAGGGTCAAGACATGCTCCAGGCTTGTGCTCCAGTAG 5660
Qy 5096 tcaaggagctggccaagatatgtggagtagtattggaacagcatttggatagacagacca 5155
Db 5661 TCAGGGACTGGCCAAATATTTGGGATGATTTGGAACTGCAATGTGCTATACAGACCCA 5720
Qy 5156 aagctcaggaagtagagaagatgaatagaacatlaagaagaccccttaccnaattgac 5215
Db 5721 AACCTCAGACAGGTAGAGAGATGATGAATGAACCATTAAGAGACCTTACCAAAATTTGAC 5780
Qy 5216 cacagagactgcatlaatgaatgaatgtgtcctcgtcccttgtgtcttttaggttag 5275
Db 5781 CACAGAGACTGGCATTAATATTTGATGTGCTGCTCCCTTGTCTTTTAAAGGTGAG 5840
Qy 5276 gaacacccctggagcagtttgggtgtgaccccttagaattgtcttaagggggaaccccc 5335
Db 5841 GAACACCCCTGGAGAGTTTGGGCTGACCCCTATGAAATTTGCTTACGGGGACCCCGCC 5900
Qy 5336 gttgacagaatctgcttgcacatagtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5395
Db 5901 GTTGGCAGAAATTTGCTTTCACATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5960
Qy 5396 taaggtcaagcgtcgaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5455
Db 5961 TAGGCTCAAGGCGCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6020
Qy 5456 ctaaggagagacttgaatttgcacatgcttccaaagtgtgagatctatgttag 5515
Db 6021 CTACAGAGAGACTTGCAGATTTCCACATGCTTCCAAAGTTGGAGATTCAAGTATGTTAG 6080
|||||
```


QY 3356 gacggagacacgagaagaactygtccagataccgcccccaacacagcaagaact 3415
DB 3353 GACGAGGACACGAGAAAGAAACTGTATGTCAGATACCGGCCCAACCCACCAAAAT 3412
QY 3416 gaggagcttttggggaacgtgattttgcagacttgatcccgagggttgacact 3475
DB 3413 GAGGAGATTGTTGGGAGACGTGGATTGTCAGACTGTGATCCCGGGGTTTGGACCTT 3472
QY 3476 agcagcccaactctacacgcttaaccaaagaagaaggagatctctctgggctccgaga 3535
DB 3473 AGCAGCCCACTACCTACCGCTAACCAAGAAAGGGAATTCCTGGGCTCCTGACGA 3532
QY 3536 ccagaaggaacttgaactgatacaaaaggccctgtagcgacactgctcggccccc 3595
DB 3533 CCAGAGGACATTTGATGCTATCAAAAGGCGCTGAGCGCCACTCTCTGGCCCTCC 3592
QY 3596 tgacgtaactaacccttacccttactatgtagtgaagcgtlaaggagtagcccgagat 3655
DB 3593 TGACGTAACTAAACCTTTACCTTTATGTGATGACGCTAAGGAGTAGCCCGGAGT 3652
QY 3656 ttaaccacaacccctagagacatgagagagacccgtgtgactactgcaaaagacttga 3715
DB 3653 TTAAACCAACCCCTAGAGACCATGAGAGACCTGTGCTACCTGTCAAAAGAGCTCGA 3712
QY 3716 tccgtgaacagtggttggcccgatgctgtagagctatgcgacgttggccatactggt 3775
DB 3713 TCCTGTACCCAGTGGTGGCCATATGCTTAAGGCTATCCAGCTGTGGCCATATGCT 3772
QY 3776 caagaagcctgaacaatgacttggagcagaataatactgtaataagcccccaatgact 3835
DB 3773 CAGAGACCTGACAAATGACTTTGGGACAGATATATACCTTAATAGCCCCCATGCAAT 3832
QY 3836 ggaagaactcgttcggcagcccccaagacagatgtagtgaacacgccccgaatga 3895
DB 3833 GGAGAAACATCGTTGGGAGCCGCCAGACGATGATGACCAACGCCCCATGACCCACTA 3892
QY 3896 tcaaaagctgcttctcaagaagaggttgaacttgcctcacacagccgttccaaacccgc 3955
DB 3893 TCAAAAGCTGCTTCTCAAGAGAGGTCACGTTCCCTCCACAGCCGCTTCAAAACCTGCG 3952
QY 3956 cactctctgcctgaagactgatacaacagtagactcatgatactgcaactattgat 4015
DB 3953 CACTCTTCTGCTGAAGAGACTGATGACCAAGTACATCATGATTCATCAACTATGAT 4012
QY 4016 tgaagagactgggttcgcgaagacacttaacagataaccgttgatggaagagtgctaac 4075
DB 4013 TGAGGAGACTGGGTCGCCAAGAGACTTACAGACATACCGCTGAGTGAAGAGTGTCTAAC 4072
QY 4076 ctgattcaactgacgaagacgatactgtgtgaagtagaagatggtctggtgcgcagat 4135
DB 4073 CTGCTTCACTGACGAGAGAGCTATGTGTGTAAGTAAGAGATGGCTGGGGCGGGGT 4132
QY 4136 ggttgaacggagaccgcgaacgacttggcgaacgactgcccgaaggaactcaagcaaaa 4195
DB 4133 GGTGAGACGGGACCCGACAGATCTGGGCGCAGCCTGCGGAGAGAACTTCAGCAAAA 4192
QY 4196 ggtcgtgactatgagccctacgaagcttggcgttgcggaaggaatccataaact 4255
DB 4193 GGTGTGACTGATGGCCCTACGCAAGCTTTGGGCTGGCGGAGGGAATTCATTAACAT 4252
QY 4256 ttaataagacagaagtagtacttggcactgacacgtacacgagggccatcatalaaca 4315
DB 4253 TTATAGGAGACAGAGTATGCCCTTTGCGACTGCACACGTACATGGGGCCATCTATAACA 4312
QY 4316 aagggggttgccttaactcaagcaggaaggaataaaagaacaaaggaatcttaagct 4375
DB 4313 AAGGGGTTGCTTACTCTGACGAGGAGGGAATAAAGAAACAAAGGGAATTTCTAAGCTT 4372
QY 4376 attagaagcttacatttccaaaagagtagactatatacactgctcggagacataga 4435
DB 4373 ATTGAAGCCGTACATTATACCAAAAAGGCTAGCTATTATACACTGTCTGAGCATCAGAA 4432

QY 4436 agccaaagatctcatatctagagggagacagatggctgacccggtgtccaaagcagcagc 4495
DB 4433 AGCTAAAGATCTCATATTCAGAGAGAAACCAAGATGGGTGACCGGGTTGCCAAGCAGGACGC 4492
QY 4496 ccagctgttaacctctgcctataatagaacgccccaaagccccagaacccagacgaca 4555
DB 4493 CCAGGGTGTAACTTCTGCTCTATAAATGCAAAATGCCCCAAAGAGGGTTAGATAA 4552
QY 4556 gtacacccctagaagacttgcgaagaagaataaagaatagaacaggttctcgtgactccga 4615
DB 4553 GTACACCTTAGAAGACTGGGAGAGATTAAGAAAGATAGACCAATTTCTGAGACATCCGGA 4612
QY 4616 ggggacctgctatacctcatatggaaggaatactctgcccccaaaagaggttgaata 4675
DB 4613 AGGAGACTGTATACCTCAAGATGGGAAGAAATCTGCCCAAGAAAGAGGTTAGATAA 4672
QY 4676 tgtccaaagatacatcgtctaaacccaactagaactaaacacctgcagcagttgtgcag 4735
DB 4673 TGTCCAAACAGATACATGCTCTAACCCACCTAGGAACTTAAACACTCTACACAGTTGCTCAG 4732
QY 4736 aacatccctctatcatgttctgagcttgcagagagtgctgactggtgtgtcaaatgtg 4795
DB 4733 AACATCCCTTATCATGTTTGTGAGCTACAGAGAGTGGCTGACTGGTGTCAACATTTG 4792
QY 4796 tgtccctgcagctgttgaatgcttaactctccagaatacctccagaaagagactaag 4855
DB 4793 TGTGCCCTGCGACCTGGTTATGCTAATCTTCACAAATGCCCTCCAGGAGAGACTAAG 4852
QY 4856 ggggaagcaccacccagcgctacacggaagtgagctcactgaggttaagcccgctaaata 4915
DB 4853 GGGAAAGCCACCCAGGCGCTCTACCTGGGAAGTGGAATCTGAGAGTAAGCCGCTTAATA 4912
QY 4916 cggaaacaaatcatctgttcttggtagacaccttccagaatgggttagaggtctaac 4975
DB 4913 CGGAACAAATFACTATTGGTTTGTAGAGACCTTTTTCAGATGGGTAGAGCTTATCC 4972
QY 4976 tactaagaagagactcaaccgltgtgtcgaagaaatactcgtgaggaatcttccag 5035
DB 4973 TACTAAGAAAGAGACTCAACCGTGGGCTTAATAAAATCTGGAAGAAATTTTTCACG 5032
QY 5036 attggaatccttaagtaataaggttcagaacaatggtccagcttctgtccaggttaag 5095
DB 5033 ATTTGGAATACCTTAAGGTAAATAGGTCAGCAATGTCACGCTTTTCTTCCAGGTAG 5092
QY 5096 tcaaggactgcccagaatattgggagattgattgaaactgcatgtgtcatatagacacca 5155
DB 5093 TCAAGGACTGGCCAAAGATTTGGGATTTGTAAGAACTGCATTTGTCATACAGACCCCA 5152
QY 5156 aagctcagagcaggtagagagatgaatagaacataaagagacccctaccaaattgac 5215
DB 5153 AAGCTCAGGACAGGTAGAGAGATGAATAGAACCATTTAAGAGACCCCTTACTAAATTGAC 5212
QY 5216 cacagagactgacataatattgactggtctctccttgccttggcttcttaaggttag 5275
DB 5213 CGCGGAGACTGGCTTAATGATTTGATAGCTCTCTGCTTGTGCTTTTAAAGGTTAG 5272
QY 5276 gaacaccccttgaacagtttggcgtgacccccctatgaattgctctaaggggagccccccc 5335
DB 5273 GAACACCCCTGACAGATTTGGGCTGACCCCTATGAATTACTTACGAGGGAGACCCCCCC 5332
QY 5336 gttgcgaagaattgcttgcataagtagtgcgtgagtgttctgttccagaccttgtctc 5395
DB 5333 ATTGGTAAATAATGCTTCTGTACATAGTGTGAGAGTGTGCTTTTCCAGGCTTTGTTCTC 5392
QY 5396 tagctcaagcgctcgaagtgtgtgagcagcagcgttgaagcagctccggaagcgcta 5455
DB 5393 TAGGCTCAAGGACACTTGAATGGGTGAGCAACGAGCGTGAAGACCTCCGGAGGCGCTA 5452
QY 5456 etc---aggagagacttgcagattccaatcgtcttccaaagtgtgagattcagctatgt 5512
DB 5453 CTCAGAGAGAGAGACTTGCAGATCCCAATCGTTTCCAACTGGAGATTCAGTCTAGT 5512
QY 5513 tagagccacacgttgcgaagaaccccgagactcgttgcgaaggaagacttatctcgtacttt 5572

Db 5513 TAAAGCCACACCGTGCAGAAACCTCGACACTCGGTGGAAGGAGGCCCTTATCTGTACTTTT 5572
Qy 5573 gaccacacaagcgtctgaaagtgcgaagaaatccccccttgatccatgcatcccaagt 5632
Db 5573 GACACACACCAAGCGGTGTGAAGTGCAGAAATCTCCACCTGGATCCATGATCCACGT 5632
Qy 5633 taagcggcgccacacccctcgcgtggtggtgagagccgaagaaagtgaatcccttaa 5692
Db 5633 TAAACGGGCGCCACCTCCGATTCGGGTGGAAGCCGAAAGAGCTGAAATATCCCTTAA 5692
Qy 5693 gcttcgcctccatcgctgtctcctactactaacaataatcccccagccagtagtaac 5752
Db 5693 GCTTGGCTTCAMCGGTGTGCTTACTCTCTCATATACTCTTAGACTAATGGTATGC 5752
Qy 5753 gectatagacagctcgaacccccatagaacttacccttaacctcgtctgattatgac 5812
Db 5753 GCATAGGAGACAGCCTGAACTCCCAATAACCTTATCTCTCCCTGTTAATTAAGTACT 5812
Qy 5813 ctgatacgggtgtcactgtaaatagactcgaagtggtgctcctcctagagcaactgtgac 5872
Db 5813 CCGGACAGAGTATTAATATCAACAACACTCAAGGGGAGCTCTTTAGGAACCTGGTGC 5872
Qy 5873 ctgaactgcattctcgcctcgcattgataaccccgctgttaaaagcaacactcccaac 5932
Db 5873 CTGATCTATACGTTTGCTCAGATCAGTTATTCCTAGTCTGA--CCTCACCCCCAGATA 5929
Qy 5933 tagtcgtagtataaggttctatctgctgcacagacagacagacagacagacagacagac 5989
Db 5933 TCCCTCAGTCACGAGATTTATGTTTCCAGGACACCAATATGGAAGAAACATTGCG 5989
Qy 5990 ggggtctcgggaatcctctcgtgtagaagtgagctgcgtacactccaaagagtagact 6049
Db 5990 GAATCCAGAGATTTCTTTTGAACATGAACTGTGAACCTCTAATATGATGATAT 6049
Qy 6050 ggaatggcgacatcctctccagacagcggtlaaattcctcctgttcaa----- 6098
Db 6050 GGAATGGCGCAACCTCAGCAGATAGATAGTTTCTTATGTCATACACCTATACCA 6109
Qy 6099 -ttccgcgcgggcaagtlacaaaatgataaactataaaagaaagtagctcccat 6157
Db 6110 GCTCGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6169
Qy 6158 cagacttagtatacctaaagtagtttact--gaaagaaacagaaataatccaa 6215
Db 6170 CAGACCTAGATTACCTAATAAATTAAGTTCACTGAGAAAGAAACAAGAAATATCTTAA 6229
Qy 6216 agtggataaagtatagctggtggaatagtttattatataagcgag--gagacag 6271
Db 6230 AATGGTAAATGATGCTTGGGAGATGATATATGAGAGCTCGGGTAAACACACAG 6289
Qy 6272 ggtccactttaacattcgccttagatagacagggagacagaaacccctgtgtgcaatg 6331
Db 6290 GCTCATTCTAATACTTGGCTCAAAATAAACACAGCTG--GAGCTCAATAGGCTATAG 6346
Qy 6332 gaccgataaagtacgtgcgaacaggggccccgcgcctgagagccacgataacttgc 6391
Db 6347 GACCAATTAAGGCTTACAGGGTCAAGACCCCAACCC----- 6385
Qy 6392 cgtgtcccaataaacctcgtcgtgcgtacataacacagcgctagcaacagtaaca 6451
Db 6386 -----AAGGACCAAGACCATCTCTTAACATTAATCTT 6415
Qy 6452 ctgagatgattcctacacagcctagaaactcccaagtgtctctgttaagacagagac 6511
Db 6416 CTGATATGAGACCCACAGTCTAACAGACGAC-----TAAATATGGGG 6460
Qy 6512 agagactctcagctcatccagagagcttccaagcaatcaactccacgcagcctgag 6571
Db 6461 CAAATCTTTTACCTCATCAGGAGGCTTTTAACTCTTAACCTCAACGACTCCAGAG 6520
Qy 6572 ccaacttctctgtgcttctactcctcaagggcctccttattagaggatagtgta 6631
Db 6520 ccaacttctctgtgcttctactcctcaagggcctccttattagaggatagtgta 6631

Db 6521 CTACCTCTTCTGTTGGCTATGCTTACCTTCGGGCCACCTTACTATGAAGATGCTA 6580
Qy 6632 aagaagaanaatcaatgtgacccaagagcaataaataatgtaactgtgggtccgaa 6691
Db 6581 GAAGAGGAAATTTAAATGTGCAAAAGAGACTTAAGACCATGACATGGGATGCCAA 6640
Qy 6692 ataagcttaccctacatgaatgttcgcggaaggaatgcatgaggaagaccccat 6751
Db 6641 ATAAGCTTACCTTACTAGAGTCTCTGAAAGACACCTCATAGAAAGGTTCCCAT 6700
Qy 6752 ccccaacaacacttctgataagtaactgtgttataagcaagcctcagaataatgalt 6811
Db 6701 CCAACCAACACTTGTGAACCACTGAACCTTAAATCAACCTCTGAGACCAATATC 6760
Qy 6812 tagtaactgttataaaggtgtggtgcatgcaactatggttaacccctgtgttcca 6871
Db 6761 TGGTACTGTGTATGACAGGTGGGCACTGTATCTGATTAACCCCTTGTTGCCA 6820
Qy 6872 cctcagcttccaacaaatccaaagattctgtgtcaltgacatccatcccgagagt 6931
Db 6821 CTTTGTGTTTAAACCAATAAGATTTTGCATTAATGTCCTCAATTTGTTCCCGAGTGT 6880
Qy 6932 actacactctgaggaagtgtcctgtgataatgactatcggtataacccaacaa 6991
Db 6881 ATTACTATCCCGAAGAAAGCAATCTGATGATGATGATGATGATGATGATGATGATGAT 6940
Qy 6992 gagaacccgataccctaccctagctgtatgtctgattaggaagcgtgtggtg 7051
Db 6941 GAGAACCCATATCTCTGACACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7000
Qy 7052 gaaacaggaagcgtgcctcctatccaggaacccaagacagctaggaaggaagcgtgtgagc 7111
Db 7001 GAACAGAGACAGCTGCCCTGT 7060
Qy 7112 taatcgcgcacatgaaagaatctccgccttaaaagatctgttagcaactagaag 7171
Db 7061 TACATGAATTTGAACAGAAAGATCTCAAGCCCTAAGAAATCTGCATTAACCTGAGAG 7120
Qy 7172 agtccctgactcttctgtcgaagtggttctcagaaacggggagattagctcgtcgt 7231
Db 7121 AATCCCTAACCTCTTATCTGAAATGCTCTACAGATTAAGAAAGGGTTGATTAATAT 7180
Qy 7232 ttctaaagaagtggtgtatgtcagccttaaaagaagatgtgtcttctatgtatagc 7291
Db 7181 TTCTAAAGAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7240
Qy 7292 actcaggaacccatcagagctccatgaacaaactagaanaaagttagagagctgaa 7351
Db 7241 ATTCAAGGSCCATCAGAGCTCATGAAACAACTTAAGAAAGGTTGAGAGAGCTCGAA 7300
Qy 7352 ggaagaagaagcgtgacaggggtgttgaaggaatgttcaagagctcctctgtatga 7411
Db 7301 GGGAAAGAAGAACTACTCAAGGCTGTTGAGAGGATGTTCAACAGGCTCTTGTGTGG 7360
Qy 7412 ccaactctctcgtcctcgtcaggggccccctagtagtctcgtcttacttaacagt 7471
Db 7361 CTACCTTACTTCTGCTTAAACAGAGACCTTAATGATGCTCTCTCTGTTACTCACAGT 7420
Qy 7472 ggcctgtcttaataatagttgtgtccttgttgaagaagtagtgaatcaga 7531
Db 7421 GGCATGATTAATTAACAAGTTAATGCTTATTAAGGAAGAAATTAAGTACAGTCCAGA 7480
Qy 7532 tcatgtacttaggcaacagtiacaaagccttctgagccaaggaatactacactagc 7591
Db 7481 TCATGTACTTAGAACAAGTACCAAGCCGCTCAG--AGGAACTGCGCGCTAGC 7537
Qy 7592 ctcccaacttcaagatlagactttaaacaagaagaatgtgggaatgaagaatga 7651
Db 7538 TCTACAGTCTTAAGATTAACTATTAACAAGAAAGAAATGAGGATTAAGAAAGATGA 7597
Qy 7652 aatgaactaacctcccaagacccaggaagttaataaagcttaaatgccccgaa 7711
Db 7598 AATACACCTTAAGTAAATGAAGAGCTTAAATTTGTTGATTCAGAGTGTGTTCTTA 7657

```

/db_xref="taxon:9823"
/cell_type="lymphocytes"
/notes="porcine endogenous retrovirus PERV-MSL"
1..564
585..2159
/codon_start=1
/product="gag protein"
/protein_id="AAC16763.1"
/db_xref="GI:3116442"
/translation="MGQVTTPLSLTLDHMTTEYKSRANHLSVQVKGKQWQTCVSEMP
PDVGVSPSGTNSSEILIAVKAVIPQTGGSHDDEPYLLTMQDLADEPPYPMWPLN
KRPQSPRIILAGEKKSNAEKVPSPHLYPLEEPPAMPDEPQSVYPLAOGAARG
PAPAPAPAVEGPAQSTRSRGATPERPTEIATPLRTGPGPLGLOPIOWPSS
ADLYNMKTNHPSESDPQGLGLVESLMKSHQTPWDCCQLLOTLETTERERILLEA
RNVPCGADGPRPRLONEIDMGPELTGPRWDYNTEBRESIKITRYMLVGLRSASRP
TNIAKREVMQGNESPSYFLERLLEAFREYEDPTSEAKSVLAATIGQSAIDIT
KRIQRLGEOEAFLRDLYAEAEKVYKRTEEERDEORKEEREERERKRROKDMIT
KILAAVEGKSNTEREDFRKIKSGPRQGNIGNRPTRDKQCAVCYCKERHGMARNCPK
KGNKGRILLAEEDKD"
2307..5744
/codon_start=1
/product="pol protein"
/protein_id="AAC16764.1"
/db_xref="GI:3133302"
/translation="MGATGQOOYPMTRRTYDVGVRVTHSEPLVPECPALYGRDL
TKMGADISEBQGEVYSANAKPIYVLTQLDDERYLSPLRYKDONQFWLBOPQAM
AETGAGLAKQVPPVYIOLKASATPVSRQYPLSKKEBQERHNOVLIOGILLVPLQ
SPMNTDLVPRKPGTNDYRVDQLEVNKRVODIHPTVLPNPMILCALPORSMTYVL
DLDADFCLLEHPTSQPLPEFEMRDGEGTGQGLTVPRLPOGKNSPTJDEALHRLD
ANRDIHQPOYTLIOYDLDLIGATKQDCLEGRKALLESDISGYSRASKAQICRE
VYTLGSLRDGOQWLTBARKKTYVPPAPTYAOMREBPSIAGFCNMLPGPATLAP
LYPLTKEGGEVNAPEHOKAFDAITKALNSAPALDPYKPTLYVDEKKVARGVL
TQTLGMRPRPVAATSKRLDPAVSQMLCKALAAVALVADAKLTGONITYIAPRA
LENIIVQRPDMAMTNAKMHQVOSLLTERVTFAPRALNAPATLLPEETDEPVNDQO
LLEEGVAKMLDIDPLTGEVLTFMTFDSYVEGKMAGAAVVDGRTIMASSLPBG
TSQKREKMLTQALRIAEKGSINIYDSSYATATAHGAIYKORLLTSAGEIKGN
KEELSLILEAVNHPKRIATIHCPGHOKAKRLISRGOMADRVAKQAQGVNLLIEM
PKAPERPRTLEDQETKIDQSETPETCTYSDGKELLPHKEGIEVOOIHRLTH
LGRKILQOLVRTSPYHVLRLPGVADSVRKHCVRQOLVANPSSRPQKRLRGSIPGHN
MEWDFTEFPAKYGKYLIVFVDTESGWVAAPYPKSTGTSVAAKLTLEETPRGIPK
VISEDNPFAEVANVSOGLALILGIDMKLCAAYPSQGOVMKRLTLELTATKTAET
GVNDMLALPEVLYPRVNTGGQELIPEYELLAGSPPVLEIASVNSADLISLQPLFE
LKALENVKORAMQOLREAVSGGGDLQIPHFOYGDVSVYRHRNAGLETWKKCYLVL
LTPPTAVKVEGISTWTHASHVKKAPPPDSQWKAETENPLKRLIHRVVPYVYNLSD"
5620..7536
/codon_start=1
/product="env protein"
/protein_id="AAC16765.1"
/db_xref="GI:3133303"
/translation="MPTTLNRHLPIRGKRPKRKILPFSASTAMFLTISTQNGM
RIGDLSNHSKPLSLTWLITDSGTGININTNGEAPLSTWMDLYLCRSVYPLSTSP
DILHAHGFVYCGPPNNGKHCQGNRPDPCQOMNCVTSNDGWPMKTSYQDQNVSYVVA
TYTSSGFNFYLTWIRTSQPKSPSDLDYLKILSEFKOEKOBILKWNMGDMWYVYGG
SGQKQGLSILTRKINOLEPMAIGPMVLTCLASGPPGQSPSSINTSGSPTESSST
TKGAKIFSLIOGAFOALNSPTPEATSSWTCILASGPPGGRKAPNTRKEHROD
CTWGSQKMLTLTEVSGKGTICGKVPRHQHCNTEAPNQTSEQYIVPGDPMACN
TGLPVCVSTLVENQTKDPCIMWQIVPARYVYBEKALITDEYDNRHNRQRETSITLAE
MLGAGVAGVOTGTAAVLVTPQOLETEGLSNLHRLVTDLDLAEKSVNSLESTLSLESE
VVLONRRGLDMLFLKEGGLCYALKKEECFYVDHGSAIRDSNNKRLERERKREKET
OGMEGPFNSIMLATTLSALTGLPLVLLLVPGPCINKLIAFERISAVQIMVL
ROQFQSPSSREARQ"
7518..7976

```

QY 61 aagacgcttcgttagatgatt-tgggtgtcgcctcttcgcagcccgagagggga 119
 DB 63 AAGACGCTTCGTAGATGATTAAAGGAGATCGCTTTCCGAGCTCGAGG----- 115
 QY 120 tttgtctttacgtcttcatttggtgtcgttgcgcggaatctctgaccacccct 179
 DB 116 --TTCTTTTGGGTCTTACATTTGGGGGCTGCTCGGAGATCTGTGGGGCCACCCCTA 173
 QY 180 aacacccagaaacacgacttggagga-aaggaatccctcttgaaagctgtg--tggtc 236
 DB 174 AACACCCGAGAACGACTTGGAGGTAAAGATCCCTTTTAAAGGTGATGATGATAC 233
 QY 237 ggcgcgcgtctctgtctgtcgtgtcgttctcgtgtgagcgccgttcgttgcagtg 296
 DB 234 GGGCGGCGTCTGTTCTGTGAGTGTCTGTTTCAAGTGTGCGCGCTTTGCTTCAAGCTG 293
 QY 297 tccctcagaacgtaagacgtgagacgtgtatcagacagacgtctgaggagatcaag 356
 DB 294 TCCTCTCAGCGCGTAAGGCTTGGGGACTGTGATACAGACGATGCTAGAGAGATCACAG 353
 QY 357 gctgcacccctggggagagcccgaggagtgaggagacagagacgctgtgtgtc 416
 DB 354 GCTGCTGCTTGGGGAGCGCCCGGAGGTGAGAGACCAAGGACGCTGTGTCTCTC 413
 QY 417 taactgtcgtcagagagacgagttcgtgtgttgaagcgaaagctccctccgcgcg 476
 DB 414 TACTGTGCTCAGAGGACCGAATTTGTTGCTGAAGCGAAAGCTTCCCTCCCGAGACCG 473
 QY 477 tccgactctttgctgtctgtgaaagacgagaggtgcgtgtgtcgtcgtatctgt 536
 DB 474 TCCGACTCTTTGCTGCTGTGGAGAGCGTGAAGCGGTGACGTTGCTGTGATCTGTG 533
 QY 537 gttctgtcgtgtgtgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 596
 DB 534 GTTCTGTTTGT 593
 QY 597 aacgtgaacacccctctgactgtacactgtgacacatttgagatagatccaggtc 656
 DB 594 ACGGTGAGACCCCTCTTAACTTACCTCTGACACTTGTGAGTGAAGTAAATCCAGGGCT 653
 QY 657 cataattgtcagttcaggttgaagaaggaacattgtcgtcgtgtgtgtgtgtgtgt 716
 DB 654 CATATTTGTGAGTTCAAGTTAAAGAGGACCTTGGCAGACTTCTGTGTCTGATATG 713
 QY 717 ccaacattcgaatgttgaatgcctcagagaggaacattcgaatcgaatcgtgtgt 776
 DB 714 CCGACATTCGATGTTGATGGCCATCAGAGGGGACCTTTAATCTGAGATTATCTGGCT 773
 QY 777 gtaagcaatcatttttcagactgagccgcgtctcatctcgtatcagagaccatattc 836
 DB 774 GTTAAAGCAGTTATTTTTCAGACCTGAGCCGCTCTATCCGATTCAGGAGCCCTATATC 833
 QY 837 cttaactgtgcagaatgttgcagaagatccctccgacatggttlaaacatgtcgaataa 896
 DB 834 CTTAACGTGGAAGATTTGGCAGAGGATCTCCGCAATGGGTAAACCATGCTGAMTAA 893
 QY 897 ccaagaagcagagtcctccgaatccgtgtctctggagagagaaacaaacactgcgcga 956
 DB 894 CCAAGAAAGCCAGGTCCCGCAATTCGTGCTTGGAGAGAAACAAACCTCGGCTGAA 953
 QY 957 aaagtgcagccctctctcgtatctaacccgagatcgaagagccgacactgtgcgga 1016
 DB 954 AAAGTCAAGGCCCTC-TCCTCATATCTAACCCGAGATTGAGAGGACCAACGGCTTGGCCGA 1012
 QY 1017 accccaactgttcccccaccccttatccagacagaggtgtcgtgtgagga--cctctgc 1075
 DB 1013 ACCCCATCTGTTCCCCACCCCTTATCTGACAGAGGTGCGCAGAGGGACCTTTTGC 1072
 QY 1076 cctcctgtgagctccggtgtgagagagactgtccgagactcggagccgagagagcgc 1135
 DB 1073 CCTCTGTGAGCTCCGGCGGTGAGGAGCTGCTGAGGGAGCTCGAGCCGAGGGGCGC 1132

QY 1136 caaccggagcggacagacagatcgacata ttaaccgttcgcacatactgcccctccat 1195
 DB 1133 CACCCCGAGGGAGACAGACGATTCGACATTTACCGTGTGCGACGTAACGCCCTCCAC 1192
 QY 1196 gccagggggccaattgcagccctccagatattgccccttctctctcgaatcctctaa 1255
 DB 1193 ACCGGGGGGCAATTTGAGCCCTCCAGATTTGGCCCTTTTCTCTGACATCTCTATTA 1252
 QY 1256 ttggaatactacatccctctctcgtggagatcccccaagcctcaggggtgtgtga 1315
 DB 1253 TTGGAAACTAACCATCCCTTTCTGGAGATCCCAACGCCCTCAGCGGTTGTGGA 1312
 QY 1316 gtccctatgtctctcacaacagcctacttggatgtatgtgtcaacagctgtgtgaagact 1375
 DB 1313 GTCCCTTATGTTCTCTACACAGCCTACTTGGATGATTTGCAACAGCTGCTGACAGACT 1372
 QY 1376 cttaacaacgagagcggagagaaattcgttagaggtcagaataaattctctcgtggc 1435
 DB 1373 CTTCAACAACGAGAGGAGAGAGAAATTTCTATTAGAGCTAAGAAATAATGTTCTGGGC 1432
 QY 1436 cgaagcgagcccaacgagttgcaaaatgagatgtgaatgtgtatccctgaatccgc 1495
 DB 1433 CGAGCGCGACCAACGCGGTTTGCAAAATGAGATTGACATGGAATTTCCCTTAACCTGCC 1492
 QY 1496 cgglttggaactacaacacgctgaagtgagagagacttgaatactatccagagctct 1555
 DB 1493 CGGTTGGAGCTACAAACGCGCTGAAGGTAGAGAGACTTGAATAATATCTGCGAGCTCT 1552
 QY 1556 ggttggtgtctccgggggcgtctcaagaagcccaactaatttgcgttaagtaagaggt 1615
 DB 1553 GGTGGCGGTCTCTCGGGGCCCTCAAGACGCCCTCAATTTGGCTAAGTAAGAGAACT 1612
 QY 1616 gatgcagggagcgaacgaacccctccglatcttcttgagaggtctatggaagcctcag 1675
 DB 1613 GATGACGAGGACCGAATGAACCCCTCTGTTTCTTGTGAGAGCTTGTGAAGCTTCAG 1672
 QY 1676 gcggttcaacccctttgatctcactcagagagcccaagaagcctcagtgccctgc 1735
 DB 1673 GCGGTACACCCCTTTGATGCCACTCAGAGAGGCCCAAAAAGCTCAGTGTGCTTGGCTT 1732
 QY 1736 catlgtggcagtggtcgtcgtgatatcagagaagaactcagagactggaaggttcaaga 1795
 DB 1733 TATAGACAGTCAAGCTTGTGATTTAAGAAAGACTTCAGACACTGGAAGGCTTACAGGA 1792
 QY 1796 ggtcgtgtacgtgtatcgtgtgagagagcagagaggtgtatcagaagaaggaacaga 1855
 DB 1793 GGTGTGTTACGTGATCTGTGATGAGAGGACCAAGAAAGTATTTCAAAAGGAGAGACA 1852
 QY 1856 agaggaagaagaaacagagaagaagaagaagaagaagaagaagcgtgtatg 1915
 DB 1853 AGAAGAAAGGAGAAACAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1912
 QY 1916 acgycagaagagaagatttgaactaagaatcttgcgcgaatgtgttgaagaagaagcag 1975
 DB 1913 ACGGCAAGAGAAAGATTGTGACTTAAGATCTTGCTCAATGTGTGAAGGAAAGCAATAC 1972
 QY 1976 ggaagagagagagatttgaagaataatgagtcagagccctgaagcagtcagaggaacctgg 2035
 DB 1973 GGAAGAGAGAGAGATTATTAGGAAATTAAGTCAAGCCCTTACAGTGAAGGAACTGTGG 2032
 QY 2036 catatgagcccccactcgcagaaagagcagtggtgtatgttaagaagaagaacactggc 2095
 DB 2033 CAATAGGACCCCACTCGACAGAGGACCAATGTGATTTGTAAGAAAGAGAGACTGGCC 2092
 QY 2096 aaggaactgcccacaagaaggaagaagaagcgaagttcctaagtaagaagaagataa 2155
 DB 2093 AAGGAACCTGCCCAAGAAAGGAAACAAAGGACCAAGGATCTTACCTTCAAGAAAGATAA 2152
 QY 2156 agatgaaggaagcgggttctcgaacccctcccgagcccaaggttaacttgaaggtgga 2215
 DB 2153 AGATTAGGGAGAGCGGGTTGCGAACCCCTCCCGAGCGCCAGGTTAATTGAAAGGTGGA 2212
 QY 2216 ggggaaaccagttgagttcctcgtgtgataccgagcggagcatcagtgctgtatacaac 2275

Db 2213 GGGGCAACCACTTGAATTCCTGGTTCATACCGGACCAACATTCATCTACTACAGCC 2272
Qy 2276 attagaanaactaaagaaaaaaatccctgggtgagatgagtgacaggaagcaagcaatg 2335
Db 2273 ATTGAGAAACTAAAGATATAAAATCTGGGTGATGGTGCCACAGGGCAACAACTAGTA 2332
Qy 2336 tccatgagactaccgaagaaccgttgatctggagtgagtgagtgagtgagtgagtgagtg 2395
Db 2333 TCCATGAGACTACCCGAAACAGTGTGAGTGGAGTGGAGGAGGAGTAAACCACTGCTTCT 2332
Qy 2396 ggtacccctgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 2455
Db 2393 GGTCACTACTAGTCCAGACACCCCTTGTAGTAGAGACTTATTGACCAAGATGGAGGC 2452
Qy 2456 tcaaatctctttgacaaagaaagcaagagtgctgtgaaatacaaaaccatcaatgtc 2515
Db 2453 ACAATTTCTTTGAAACAGGAAACAGAGTGTCTGCAATTAACAAACCTATCACTGT 2512
Qy 2516 gttgacccctcaatagatgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 2575
Db 2513 GTTGACCTCCATTTAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 2572
Qy 2576 agatatacagtcctggttgagcaagttccccaagcctgggacagaaacaggaagtgag 2635
Db 2573 AAATATACATTTCTGTTGGAACAGTTTCCCAAGCCTGGGACAGAACCGCAGGGATGG 2632
Qy 2636 ttggcaagcaagatgccccaagtgatatacactgaagcagtgctacacagatc 2695
Db 2633 TTTGGCAAGCAAGTTCCTCCCAACAGATTTCACTGAGAGCCAGTGGCACACAGTCTC 2692
Qy 2696 agtcagcaagtgaccccttgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 2755
Db 2693 ACTCAGACAGTACCCCTTGAGTAAAGAGCTCAGAGAGGATTCGGCCGCAATGTCAGAG 2752
Qy 2756 attaaaccaagggagcatcctagtcctgtcccaatcccttggaactcctccctgtaacc 2815
Db 2753 ATTATTCACAGGAGGATCTGATGTCCTGTCATCTCCCTGAGATCTCCCTGCTACG 2812
Qy 2816 ggttgaagagcctggagcaaatgatatccagtgagtgagtgagtgagtgagtgagtgag 2875
Db 2813 GGTGAGAAAGCCTGGAGACTTAATGATGATGACAGTACAGGACTTGAAGAGGTAAATTA 2872
Qy 2876 aaggttgacagcatcacacccaaggtccccaagccttataaccccttgagcgcctccc 2935
Db 2873 ACGGGTGAGATATACCCCAACAGTCCCAACCCCTTATACCTCTGTGTGCTCTCC 2932
Qy 2936 gcttgaacggaactggtacacagtgatlgagtgacttaaaagatgctctctctgctgagat 2995
Db 2933 ACCCCACGAGAGCTGGTATACAGTATGAGATTAAAGATGCTTCTCTGCTGAGATT 2992
Qy 2996 acaaccactagccaacacacttttgctctgagtgagtgagtgagtgagtgagtgagtgag 3055
Db 2993 ACACCCCACTAGCAACCACTTTTGGCTTCCAAATGGAAGATCCAGATACCGGAGAGAAC 3052
Qy 3056 cggagcagctcacttgagcagcagctgcccgaaggttcaagaactccccaacatcttga 3115
Db 3053 CCGGCAAGCTCAGCTGAGACCCGAGCTGCCCAAGGTTCAAGAACTCCCGACATCTTTGA 3112
Qy 3116 cgaagcctcactacaggaagccttgagcgaacttcagatcccaacccctcagtgagtgag 3175
Db 3113 CGAAGCCCTACACAGAGAGCTGCCCAACTTCAGGATCCCAACCCCTCAGGTGACCTCT 3172
Qy 3176 ccaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 3235
Db 3173 CCAAGTACGTGATGAGCTCTTCTGGCGGAGCCACCAACAGAGATGCTTATGAGAGGAC 3232
Qy 3236 gaagcagctactgctgagatgtctgaactagtgacagagcctctgcttaaaagagcca 3295
Db 3233 GAAGGCACTACTGCTGGAATTTGCTGAGCTAGGCTACAGAGCTCTGCTGCTAAGAGGCCA 3292
Qy 3296 gatttgaggaagagtgatatacactctgggtgagtgagtgagtgagtgagtgagtgag 3355
Db 3293 GATTTCAGAGAGAGATTAACATTTCTGGGTGATACCTTTACGGGACGGGACGATGGCT 3352
Qy 3356 gacagaggaagcaaggaagaaacttgatccagataccgcccacacacagccaacagtg 3415
Db 3353 GACGAGAGCAGGAGAAANACTTACTCAGATACCGGGCCCCACAGCCAAACAAAT 3412
Qy 3416 gaagagtgcttggagcagctgagatcttgacagctgagtgagtgagtgagtgagtgagtg 3475
Db 3413 GAGAGAGTTTGGGACAGAGCTGATTTTGACAGACTGTGAGATCCCGGGGTGGGACCTT 3472
Qy 3476 agcagccccactacccgttaacaaagaagggagtgctctgggtgagtgagtgagtgag 3535
Db 3473 AGCAGCCCACTCTACCCCTTAACCAAGAAAGAGGAGATTTCTGGGCTCTGAGCA 3532
Qy 3536 ccagaaggaacttgatgactaacaagaagcctgctgagcagcagcagcagcagcagcagc 3595
Db 3533 CCAAGAGCACTTTGATGCTATCAAAAAGGCCCTGCTGAGGCACTGCTGCTGCTGCC 3592
Qy 3596 tgaagtaactaaaccccttaaccccttaagtgtgagtgagtgagtgagtgagtgagtgag 3655
Db 3593 TGACGTAACTAAACCTTTACCTTATGTGATGAGCGTAAAGGAGTAAAGCGGAGAGT 3652
Qy 3656 tttaacccaacccctgagcagcagtgagtgagtgagtgagtgagtgagtgagtgagtgag 3715
Db 3653 TTTAAACCAACCCCTAGAGACATGAGAGAACCTGTGTGCTTACCTGTCAAAAGAGCTCGA 3712
Qy 3716 tccgtgagcagtgatgagcagtgatgctgagtgagtgagtgagtgagtgagtgagtgag 3775
Db 3713 TCCGTGAGCCAGTGTTGGGCCATATGCTGAAAGGCTATGCAAGCTGTGGCCATACGT 3772
Qy 3776 caagagcagctgacaaacttgactctggagcaagaataactgtatagcccccaatgag 3835
Db 3773 CAAGGAGCTGACAAATTTGACTTTGGAGCAAGATATTAACGTATATACCCCTCATGAT 3832
Qy 3836 ggaagaatcgttgaggaagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3895
Db 3833 GGAGAAATGCTGTGGAGCCCCCAGCCAGCTGATGATACCAAGCCCGCATGACCACTA 3892
Qy 3896 tcaagcctgctctcacaagaagagtgactctgctcacaacagcagcagcagcagcagc 3955
Db 3893 TCAGAGCTGCTTTCACAGAGAGGCTGACGTTGCTCCACAGCCGCTCTCAACCTGCG 3952
Qy 3956 caactctctgctgaggaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 4015
Db 3953 CACTCTTCTGCTTAAGAGACTGATGAACCAAGTACATGATGATGATGATGATGATGAT 4012
Qy 4016 tgaaggaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 4075
Db 4013 TGAGGAGACTGGGTCCGCAAGGACTTTACAGACATACCGCTGACTGGAGAGTGTAAAC 4072
Qy 4076 ctggtcactgaaggaagcaactatgctgtgagtgagtgagtgagtgagtgagtgagtgag 4135
Db 4073 CTGGTTCATGAGCGGAAGCACTATGTGTGGAAGGTAAAGGATGCTGGGGGGGGCT 4132
Qy 4136 ggtgagcaggaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4195
Db 4133 GGTGAGAGCGGACCCGACAGATCTGGGACAGAGCTCCCGGAGAACTTCAGACAAA 4192
Qy 4196 ggtgagcagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 4255
Db 4193 GGTGAGAGCTATGAGGCTTACAGCAAGCTTTTCGGCTGGCCGAGGAAATTCATTAACAT 4252
Qy 4256 ttatacgaagcagcagtgatgctcttgagtgagtgagtgagtgagtgagtgagtgagtgag 4315
Db 4253 TTATATCGACAGCAGGTATGCTTGTGAGCTGACAGGTATATGAGGCTCATCTATTAACA 4312
Qy 4316 aaggggttgcttactcagcagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 4375
Db 4313 AAGGGGTTGCTTACCTCAGCAGGAGGAAATTAAGAGCAAAAGGAAATTTCTAAAGCT 4372
Qy 4376 attaaagccttaactctgccaagaaagtgactatatacagtgctctgagatcagaa 4435
Db 4373 ATTGAAGCCGTACTATTACCAAAAAGGCTAGCTATTATATACACTGCTGAGATCGAA 4432

4436 agccaaagatcctcatalctagagggaaccagatggtcgaacgggttgcgaacgaagcagc 4495
4433 AGCTAAAGATCTCATATCCAGAGAAACCATGGCTGACCGGGTTCCCAACAGACGACC 4492
4496 ccaaggtgttaacctctcctcctcctcctcctcctcctcctcctcctcctcctcctc 4555
4493 CCAGGCTGTAACT 4552
4556 gtacacccctagagactggaagataaagaatagaacggtctctcctcctcctcctcctc 4615
4553 GTACACCCCTAGAGACTGGCAAGATATAAAGATAGACAGTCTCTCTGACACTCCGGA 4612
4616 gggagactgctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 4675
4613 AGGAGACTCTCTACCTCAGATGGAGAAATCTCTGCCCAACAAGAGGTTGATTA 4672
4676 tctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 4735
4673 TCTCCAAAGATACATCT 4732
4736 aacatccctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 4795
4733 AACATCCCTCTATCATCT 4792
4796 tctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 4855
4793 TCTGCT 4852
4856 gggagacccacccacccacccacccacccacccacccacccacccacccacccacccac 4915
4853 GGGAGACCCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCCAC 4912
4916 cggaaacaaatcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 4975
4913 CGGAACAAATACCTATCT 4972
4976 tactaagaagaagactcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 5035
4973 TACTAAGAAGAAGACT 5032
5036 atttgaatacctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 5095
5033 ATTGGAATACCT 5092
5096 tcaaggactggtcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 5155
5093 TCAGGACTGCTGCAAGATCT 5152
5156 aagctcaggaagaagtagagagatgaaatagaacattagaagacccctcctcctcctcct 5215
5153 AAGCTCAGGACAGTAGAGAGATGAATAGAACATTAAAGAGACCTTCTCTCTCTCTCTCT 5212
5216 caccagaaactgctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 5275
5213 CGCGAGACTGCTGCT 5272
5276 gaaacacccctggaagactggtcctcctcctcctcctcctcctcctcctcctcctcctc 5335
5273 GAACACCCCTGGACAGTTTGGGCTGACCCCTCTATGATCTCTCTCTCTCTCTCTCTCTCT 5332
5336 gttgcaagaatgctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 5395
5333 ATTGCTAGAAATCT 5392
5396 taagctcaagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 5455
5393 TAGGCTCAAGGCACTTGAAGTGGGAGACACACAGCTGAGGCACTCCGGAGGCTCTA 5452
5456 ctc---aggaagagacttgaagttccacatcgtctcccaagttgagatcagctatgt 5512
5453 CTCAGGAGGAGAGACTTGCAGATCCACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5512

5513 tagagccacacgttcaggaacccctcctcctcctcctcctcctcctcctcctcctcctcct 5572
5513 TAGAGCCACCCCTGAGAGAAACCTGAGACTCGGTGGAGGCGCTTATCTCTCTCTCTCTCT 5572
5573 gaccacacacacgttcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 5632
5573 GACCACACACACGCTGTGAAGTGGGAGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5632
5633 taagcggcggcgcacccctcctcctcctcctcctcctcctcctcctcctcctcctcctc 5692
5633 TAAACCGCGCGCACCT 5692
5693 gttcgcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 5752
5693 GCTTCGCT 5752
5753 gctctatagaagcgtcgaacccctcctcctcctcctcctcctcctcctcctcctcctcctc 5812
5753 GCATAGAGAGACAGCTGACCTCCATTAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5812
5813 ctgatacgggtgtcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 5872
5813 CCGGACAGGATTAATATATCAACACACTCAAGGGAGGCTCTCTCTCTCTCTCTCTCTCTCT 5872
5873 ctgaactgcatctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 5932
5873 CTGATCTATACGTTTGGCT 5929
5933 taagcgtatgtatggtcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 5989
5930 TCTCTCATGCTCAAGGATTTATGTTGCTCCAGACACCAATTAATGGAACATTTGGG 5989
5990 ggggtctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 6049
5990 GAATCTCCAGAGATTTCT 6049
6050 ggaatcggcggcgtcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 6098
6050 GGAATCGGCCACCT 6109
6099 -ctcggccggcggcgtcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 6157
6110 GCTCTGACAAATTTATTTCT 6159
6158 caagctatgatactcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 6215
6170 CAACCTAGATTTCT 6229
6216 agtgaataaagtgaagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 6271
6230 AATGGSTAAATGGTATGCTTGGGAGATGTTATTTATGAGAGCTCGGGTAAACACAG 6289
6272 gttcgcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 6331
6290 GCT 6346
6332 gacccgataaagctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 6391
6347 GACCAATACGCTCTTGAAGGCTCAAGACCCCAACCC----- 6385
6392 cgggtcccaatlaacctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 6451
6386 -----AAGGACCAAGACCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 6415
6452 ctggtatgtatctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 6511
6416 CTGATCAAGACCCCACTGATCTTACGACACAGAC-----TAAATGGGGG 6460
6512 agagactcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 6571
6461 CAATACTTTTATGCT 6520
6572 ccacttctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 6631

Db 6521 CTACCTCTTCTTGGCTATGCTTGGGCGCCACTTATCATTAAGAGATGGCTTA 6580
 Qy 6632 aagaagaanaatcaatgtyaccaagaagacatagaatacaatgtaacgagcgcga 6691
 Db 6581 GAAGAGGGAATTCATGTGCACAAAGAACATAGAGACCAATGACATGGGGATCCCAA 6640
 Qy 6692 ataagcttaccctcactgaagttccggaagaggagacatgcatagaaagctccccc 6751
 Db 6641 ATAGGCTTACCTTACTGAGGTTCTGAAAAGGACCTGACATGAGAAAGTTCCCCAT 6700
 Qy 6752 cccacaacaccccttgctatagtaactgctgttaagagagcctagaanaacatgatt 6811
 Db 6701 CCCACCAACACCTTTGTAACACATGAGACCTTATATCAAACTCTGAGAGCAATATC 6760
 Qy 6812 tagtaacctgtataacagtggtgagcatgcaatactggtgtaacccctggttccca 6871
 Db 6761 TGGTACCTGGTWTGACAGGTGGGGCATGTATATCTGATTAACCCCTTGTGTTCGA 6820
 Qy 6872 cctcagcttcaacacatccaaagatttctgctgtaactgtaacatcgtcccccagtg 6931
 Db 6821 CTTTGCTTTTAAACCAAACTAAAGATTTTTCATATAGGTCATAATGTTCCCGAGTGT 6880
 Qy 6932 actaccatcctggaagtgctccttgatgaataactacatcgtgtaacccacaanaa 6991
 Db 6881 ATTACTATCCCGAAAAGCAATCCTTGATGAATATGACTACAGAAATCATCGACAAAGA 6940
 Qy 6992 gagaaaccgtaacccttaaccatgtaactgctgtaactgagagcgcgttgctgag 7051
 Db 6941 GAGAACCCATATCTGACACTGCTGTGATGCTGAGACTGGAGTGGGAGCGAGGTGTG 7000
 Qy 7052 gaacagagacagctgcccctatacagagagacagacagcctagaagaagacatggtgag 7111
 Db 7001 GAAGAGGAACAGCTGCCCTGTGTCAGGACACAGAGCTAGAAACAGACTTATAGTAAAC 7060
 Qy 7112 tacatgagccatgagaagaatctccgagccttaagagagctgtgtaagaactgaag 7171
 Db 7061 TACATGCAATTTGTAACGAAAGATCTCCAGCCCTAGAAAATCTGTACATGACTGGAG 7120
 Qy 7172 agtccctgacttctgtctgaagtggtctcaagaacggagggatlagactgctgt 7231
 Db 7121 AATCCCTAACCTCTTATCTGAAAGTACTCTTACAGAAATGAAGGTTATGATTTATAT 7180
 Qy 7232 ttccaagaagagtggtgtgtgcaagccttaagaagaagtgtcttctatgagatc 7291
 Db 7181 TTCTAAAGAAAGAGGATTTATGTGTAGCTTGAAGAGGAATGCTGTTTATGTGATC 7240
 Qy 7292 actaagagacatcagagactcatalgacaagaacttiagaanaaagttagaagagcgtcga 7351
 Db 7241 ATTCAAGGGCCATCAGAGACTCCATGAACAAGCTTACAGAAAGCTTGAGAGACCTCGCA 7300
 Qy 7352 gggagaagaagagctgaccaggggtgtgttgaagaagtgttcaacaggtctccctgata 7411
 Db 7301 GCGAAAAGGAATACATCAAGGGGTGTTTGAAGGATGTTCAACAGGTCCTTTGTGTGG 7360
 Qy 7412 ccaccctgcttctgctcgaaggggcccctagtagctgcctgccttcttactaagctg 7471
 Db 7361 CTACCTACTTCTGCTTTAAACAGAGACCTTAATAGTCCCTCCTCCTTACTCAGAGTGG 7420
 Qy 7472 ggcctgtctaataatagagttgtgtgcttgttgaagaacagagtgagtgcaatcaga 7531
 Db 7421 GGCAGATGATATTAATCAAGATTAATGCTTCAATTAAGCAACAAATGAAGCGATCCAGA 7480
 Qy 7532 tcatgagacttagcaacagtagcaagagccttcgagccaagaagaaactgactcagc 7591
 Db 7481 TCATGTAAGTCTTAGACAACAGTACCAAGCCGCTTAC---AGGAGAGCTGGCCGCTAGC 7537
 Qy 7592 ctccacagttctaagatlagaaactatacaagaacaagaagtgaggatgaagaagatga 7651
 Db 7538 TCTACCAAGTCTTAGATTAAGAACTATTAAACAAGAGAAAGTGGGGAATGAAGAGATGA 7597
 Qy 7652 aatgcaactaaccctccagaaccaggaagttaataaagaagctctaattgccccgaa 7711

Db 7598 AATACAACTTAAGCTTAATAGAAAGCTTAATAATGTTCTGAATTCAGAGTTGTCTTA 7657
 Qy 7712 ttacagaccctgctgctgcccagtaaatagtagaaggtcaactctccatgttccag 7771
 Db 7658 TAGGTAAAGATTAAGTTTGTCTTTTCTGTTTAAATATCGGAAGTAATAATAGCCCTGAG 7717
 Qy 7772 gcttgctatccctgctc-aaglaagaatacaggaatgagttgac---taagcttat 7826
 Db 7718 TACATGCTCTTAAGCATGAACCTTCTGAACTATTGAGATTAACAAAGAAAGGAGTTT 7777
 Qy 7827 ctggaattctglaaacaactgagcagcatagagaatgattatgatacatgtgacgacctag 7886
 Db 7778 CTAACTGCTTGTGTTAGCTTGTAAACCTGTTGCGCCATAAAGATGTTGAATGTGAT 7837
 Qy 7887 tgacctatctcaactcagcaatcgtgtaactctgcccagagagccacagatgagacctcc 7946
 Db 7838 ACACATATCTTGTGTACCAACATGTCTCCCAACCCCAACATGCGCAAAATGTGAATCTC 7897
 Qy 7947 ggaactatltlaaaatgattgtgtccacagagagcgagctctcgatatttaaatgattg 8006
 Db 7898 TAAACAAATTAATTAATTTGTCACAGAGCGCGCTCTGAGATTTTAAATGACTG 7957
 Qy 8007 gtccatgagagcgagcctcctcgatatlttaaatgattgtgtgtgacgacagccttg 8066
 Db 7958 GTTGTG-----GATATTTTGAATATGTTGTGAAGCGCGGCTTGTG 8002
 Qy 8067 ttgtgaaccccaataaagctgtcccgatctcgacatcgagcgagcgcagctctacacctg 8126
 Db 8003 TTGTGAACCCCAATAAAGCTGTCCCGACTCCACACTCGGGGGCGGAGTCTTACCCCTG 8062
 Qy 8127 cgtgtgtgacagctgtggt 8186
 Db 8063 CGTGGTGTACGACTGTGGGCGCCAGCGGCTGTGAATTAATAATCTCTGTGCTTGTGAT 8122
 Qy 8187 caaaaaaaa 8196
 Db 8123 CAAAAAANA 8132

RESULT 11
 AF038599
 LOCUS AF038599 8132 bp mRNA MAM 21-MAY-1998
 DEFINITION Sus scrofa porcine endogenous retrovirus tsukuba-1 mRNA, complete sequence.
 ACCESSION AF038599
 VERSION AF038599.1 GI:3133300
 KEYWORDS
 SOURCE
 ORGANISM
 Sus. scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 1 (bases 1 to 8132)
 AUTHORS Akiyoshi, P.E., Denaro, M., Zhu, H., Greenstein, J.L., Banerjee, P. and Fishman, J.A.
 TITLE Identification of a full-length cDNA for an endogenous retrovirus of miniature swine
 JOURNAL J. Virol. 72 (5), 4503-4507 (1998)
 MEDLINE 98216827
 REFERENCE 2 (bases 1 to 8132)
 AUTHORS Fishman, J.A.
 TITLE Direct Submision
 JOURNAL Submitted (16-DEC-1997) Infectious Disease Unit, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02114, USA
 3 (bases 1 to 8132)
 REFERENCE Fishman, J.A.
 TITLE Direct Submission
 JOURNAL Submitted (15-MAY-1998) Infectious Disease Unit, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02114, USA
 REMARK
 COMMENT Sequence update by submitter
 On May 16, 1998 this sequence version replaced gi:3116437.
 FEATURES
 location/Qualifiers
 1..8132
 /organism="Sus scrofa"

Oy	297	tactctcagaccgttaagagacctcgagagactcgtatcgaagaacgtgctagaagatcaag	356
Db	302	tctctctcagcccgtaaggccctggggggacttgtaacacacgctgcgtaggaggtccacg	361
Oy	357	gctgcacacccttgaggagaccgccggaggatctggaggagacgaaggaacgtctgtctcc	416
Db	362	gctgctgcctcctgggggagaccgccggaggatgtagagagacgagggagaccctggtgtctcc	421
Oy	417	tactctcgtatcaagagacccaatctcttattttaaagcaaaaacgttccctctccacggca	476

Db	482	TCCGACCTCTTTTGGCTGTTTGTGGAATACGCGAGACGGGTACGCGTCTTCGATCTCTTG	5411
Oy	537	gtttctgtctcgtgtgtgtcttgtcttgtgcgtctgtcttaacagctttaalatygaacag	556
Db	542	GTTCCTGTTTGTGTGCTTGTGTCCTTCTTACAGTTTAAATATGAGGACAG	601
Oy	597	acagttgaactaccccccttagtttgaactctcgacccattgagactaaagttagatccagaagct	656
Db	602	ACGGTAGACGACCCCTCTTAGTTGACTCTCGACCATTTGAGACTGAAGTAAATCCAGGGCT	661

[illegible]

OY	717	ccaacatcgcgtglttggtagtgccalcagagggaaccttaattctgaataatcctcgct	776
Db	722	CCGACATTCGATTGTGGATGGCCATCAGAGGGGCCCTTAAITCTGAGATTATTCGGGCT	781
OY	777	gttaagccaatcatlltttcagacttgcaccggtcctcatctcgtatcaagagccctatatc	836
Db	782	GTTAAACCAATTATTTTTTCAGACTGCACC GGCGCTCATCCGATCAGAGGCCCTTATTC	841
OY	837	cttacgtgcagaagatttggcagagaagatcctccgcgcatalggtltaaaccatlgctaaataa	896
Db	842	CTTACGTGGCAAAGATTGGCAGAGAGGATCTCCGCCATGGGTTAACCATTGGCTGATTAAG	901
OY	897	ccaagaagccaggtccccgaatcctcgtcctcttgagagaaaaacaacactcgcgcga	956
Db	902	CCAAGAAGGCCAGGTCCCCCAATTCTGGGCTCTTGAGAGAAAAACAACACTCTGGCTGA	961

QY 957 aaagtcgagccctcttcctcgtatctacccgagatcgagagccgcgcgacttggccgga 101
|||||
962 AAAGTCGAGCCCTCTTCCTCGTATCTACCCGAGATCGAGAGCCGCgcgacttggccgga 102
|||||

101/ accccaacatcgttccccacacccctatccagcacagggctgcctgtagggga -cctctgc 107
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1021 ACCCCAATCTGTTCCCCACACCCCTTATCTGACACAGGGTGCCGCGAGGGGACCCCTTTCG 108

QY	1136	caccccggaacgagacagacagatctcgatattacgcctcgagacattatggccctcccat	1191
Db	1081	cctctctcggagctccggcggttggaaggagacctttgttcaggagacatcggagagccgagagccgc	1144
QY	1141	caccccggaacgagacagacagatctcgatattacgcctcgagacattatggccctcccat	1200
Db	1196	gccagggggcgaatttgagagccctccagattatggccctttctctctgcagatctctataa	1251
QY	1201	acgggggggccaatttcagagccctccagattatggccctttctctgcagatctctataa	1266
Db	1256	ttgaaacatcaacatcccccctctctcggagagatccccaagcgtcaacagggttgttga	1311
QY	1261	ttggaaaacatcaacatcccccctctctcggagagatccccaagcgtctacggggttgatga	1322
Db	1316	gtcccttatgtctctcaccagagcctacttggatgatgtgtcacaacagctgtctgacaacat	1377
QY	1321	gtcccttatgtctctcaccagagcctacttggatgatgtgtcacaacagcctctgacaacat	1380

Db 5758 GCGATAGAGACAGCCGGAACCTCCATTAACCCCTTATCTCTACCTGGTTAATTACTGA 5817
Oy 5811 cccatatacgggtgtcaactgttaataaatacactcagagtgctgactcagagcactgtg 5870
Db 5818 CTCGGGACAGAGTATTATATCAACACACTCAAGGGAGGCTCTCTTTAGAACCTGTGTG 5877
Oy 5871 gacctgaactgtatctgctcgcagatgtatlaaccgcgctgtttaaagaacactccca 5930
Db 5878 GCGGATCTATAGCTTTGGCTCAGATCATGTAATTCCTAGTCTGAC---CTCACCCCGAGA 5934
Oy 5931 cctagtcgtatgtataggtgtctatctgctcgcagagcaag---agaagaagaatctg 5987
Db 5935 TATCTCATCTCAGCGATTTTATGTTGGCCAGGACACCAATATATGAAAAACATTG 5994
Oy 5988 tggaggtctctggagatccctctgtaagaga tggagctgcgtcaactcccaaga tggaga 6047
Db 5995 CGGAATCCAGAGATTTCTTTTGTAAACATGGAATGTAACCTTAATGATGATA 6054
Oy 6048 ctgaaatggccgactctctccagagccggttaaattcccttltcaatccgccc 6107
Db 6055 TTGGAAATGGCCACCTCTCAGCAGATAGGTTAAGTTTCTTATGTCAACACCTATAC 6114
Oy 6108 gggc-----aagtaaaaaatgataaactataaagaataagagctgtcccc 6155
Db 6115 CAGCTCTGACATTTAATTATCTGACCTGGATTAGAACCTGGAAGCCCAAGTGTCTCC 6174
Oy 6156 atcagactagatatactaaagataagtltaact--gaagaagaacagaaatatca 6213
Db 6175 TTCAACCTAGATTAACATAAAATTAAGTTTCACTGAGAAAGAAAGAAATATCT 6234
Oy 6214 aaagtgagtaagtgatagagtgaggagatgtttatataatagggg---ggagc 6269
Db 6235 AAAATGGGTAATGGATGTCTTGGGAGATGTTATATTATGAGAGCTGGGTAAACAC 6294
Oy 6270 aggttcacttaaacatctgcctcagatagagaagggagcagaaacccctgtggcaat 6329
Db 6295 AGGCTCCATTCTTACATATCGCTCAAAATAAACAGCTG---GAGCTCCCAATGTCTAT 6351
Oy 6330 gggagccgaataaactgtgctgtaagaagggcccccggcctgtgagccacgtataact 6389
Db 6352 AGGACCAATFACGGCTTGACGGGTCAAGAACCCCAACCC----- 6392
Oy 6390 ggcgggcccacttaactcgtcgtcgccagacataacagccgcctagcaacagtaac 6449
Db 6393 -----AAGSACAGSACCATCTCTTAACATAAC 6420
Oy 6450 caatgattgattcctaccacaagcctagaactcccaagtgcttctgttaagacag 6509
Db 6421 TTCTGTGATCAGACCCACCTAGCTAGCTAGCAGCA-----CGACTAAATATGGG 6465
Oy 6510 acagagactcctcagctcctacccagggagcttccaagcactcaactccagccagc 6569
Db 6466 GGGCAAACTTTTATGCTTATCCAGGAGCTTTTCAAGCTTTAATCTCCACGACTCCAA 6525
Oy 6570 tggcactctctctgtgtgtctgtctacccagggcctccttaltataggggagtg 6629
Db 6526 GGGTACTCTTCTTGTGTGGCTATGCTTATGCTTTGGGCCACCTTACTATGAAGATGGC 6585
Oy 6630 taaagaagaanaattcaatgtgacaaagacatagaatcaatgtacatgtggggtccg 6689
Db 6586 TAGAAGAGGGAATTCATATGTCAGAAAGAACATAGAGCAATGACATGGGAGATCCA 6645
Oy 6690 aaataagcttacccctcaggaagtctcggaagagagatcatagagaagaagctcccc 6749
Db 6646 AAATTAAGCTTACCTTACTGAGGTTTCTGGAAGAGGACCTGATAGGAAGGTTCCGCC 6705
Oy 6750 atcccaacaacacttgcctatagtaactgtgttla tggacagggcctcagaaatcaat 6809
Db 6706 ATCCCAACCAACCTTTGTATACCACTGAAAGCTTTAATCAAACTCTGAAAGATCAATA 6765
Oy 6810 tttaagtaactgttataacaggtgtgtggagatgcaataactgtgttaacccctgtgttc 6869
Db 6766 TCTGTACTCTGTTATGACAGGAGTGGGCAATGTAATACCTGATTAACCCCTGTGTTC 6825
Oy 6870 caactagcttcaaccaatccaagaattcgtgtgcatcgtgcccgaatcccgagct 6929
Db 6826 CACCTTGGTTTAAACCAATTAAGATTTTGGATTATAGGTCACAAATGTTCCCGAGT 6885
Oy 6930 gtaactaacctcgaagagtggtccttgaataataatgaactcgttataaccca 6989
Db 6886 GTATTACTATCCGAAAGAACATCTCTGATGATGATGATGATGATGATGATGATGAT 6945
Oy 6990 aagaagaacccatcccttaacccctgaactgaatcgtgaatcgtgaatcgtgaatcgt 7049
Db 6946 GAGAGAACCCATATCTCTGACACTGCTGATGAGCTGGAAGTGGAGTGGAGCAAGTGT 7005
Oy 7050 aggaacaagagacgctgcccctgaatcagagacagacagcctgaagaagaagactgtgta 7109
Db 7006 AGGAACGAGAACGCTGCTGCTGACGAGGACACAGCACTGGAACAGACTTATGTA 7065
Oy 7110 gtlacatgcgcaatgaagaagatcctcagaccttaagaagagctgttagcaactaga 7169
Db 7066 CTAACATGCAATTTGTAAAGAGATCTCCAAAGCCCTAGAAATACTGTAGTAACCTGGA 7125
Oy 7170 agagtcctgaactcttctgtcgaagtgttcaagaaacggagggatagatctgct 7229
Db 7126 GGAATCCCTAACCTCTTATCTGAAGATGCTTACAGAAATAGAGAGGTTATATTAT 7185
Oy 7230 gtttcaagaagaagtggtgtatcgtgcagccttaagaagaatgtgtctatgtaga 7289
Db 7186 ATTTCTTAAGAGAGAGATTTATGTGTAGCTTGAAGAGAGATGCTTTATGTGGA 7245
Oy 7290 tcaactagaagcaatcagaactcagaaactgaagaactgaagaagaatgaagaagtgctg 7349
Db 7246 TCATTGAGGGGCCATCAGAGACTCCATGAACCAACTTGAAGAAAGTGAAGAGCTCG 7305
Oy 7350 aaggaaagaagagctgacagaggtgtgttgaagaatgttcaacagctcctgtgat 7409
Db 7306 AAGGAAAGAGAACTACTCAAGGCTGTTGAGGATGTTCAACAGCTCTCTTGTT 7365
Oy 7410 gaccacactgtcttctgtcgtcgaagggccctagtagtctgtcctgttaactaagt 7469
Db 7366 GGCATACCTATTCTCTGCTTAAACAGAGACCTTAAATGATGCTCTCTGTATACAG 7425
Oy 7470 tggcctgtcttaataatagttgtgtcctgttgaagaagagtggtgagcca 7529
Db 7426 TGGCCATGTAATTTAACAAGTTAATGCTTATTAAGAAAGAAATGATGAGTCA 7485
Oy 7530 gataatgtaacttaggagacagatcaagaagcctctgagcaagaagaactgactcta 7589
Db 7486 GATCATGTGTTATGACACAGCTACCAAGCCCTCTAGC---AGGAGAGCTGGCGCTA 7542
Oy 7590 gcttccagcttcaatagatlaagaactlaacaagaagaagtggggaatgaagaatg 7649
Db 7543 GCTTACACAGTTTCAAGTTAGACTATTATCAAGAGAGAAAGTGGGAATGAAGATG 7602
Oy 7650 aaatagaaactaacctccagaaacccaagaatlaataaagctctaagtgtccccg 7709
Db 7603 AAAATTAACCTTAAGCTTAATGAAGAGCTTAAATTTGTTCTGAATTTCCAGACTTGTCT 7662
Oy 7710 aattacagacactgtcgtccagtaaatagtagaaggtcaactcctatgttca 7769
Db 7663 TATAGTAAAGATTTAGTGTTTTGTCTTTTAAATATGCGGAAGTAAATAGGCCTG 7722
Oy 7770 gggcctgtactcctgctc--aagtaagaatacagaagaatgagtgac-----taatgct 7824
Db 7723 AGTACATGTCTAGAGATGAAACTTCTGGAACATTTTGAATTAACGAAGAAAGGAGT 7782
Oy 7825 atctgtatctgttaaaactcgtcagcactagaagaatgtattacacatgacagccct 7884
Db 7783 TTCTAAGCTGTGTTTACTCTTGTAAACTGTTGCCCTATTAAGATGTTGAATGTG 7842
Oy 7885 atgacatactcaactcgaactgtcactcgtccagagagccacagcagatgcygact 7944
Db 7843 ATACACATATCTTGTGTACACAGATGTCTCCCAACCCGAAAGATGCGCAATGTATAC 7902

OY	7945	ccggagcgattttaaataatgatttgctccagcgagacgggggctctcgaataattttaaaatgat	8007
Db	7903	TCTAAACAATTTAAATTAATTAATGGCTCCACCAACGGGGGCTCTCGAAGTTTTAAATTTGAC	7962
OY	8005	tgtctcatcgtgagcgcggcgctctcgataattttaaatcgaattgcttgtagacacagcgtt	8064
Db	7963	TGGTTTGT-----GATATTTGAAATGATTTGGTTGTAAAGCGGGGCTT	8007
OY	8065	tgtttgtaaaccccaataaactgtctccgaattccgaacttcggggccgcgcagctctataccc	8124
Db	8008	TGCTGTCAACCCCTTAACAGCTGTCCCGAATCCACACTGGGGCGCGCACTCTCAACCCC	8067
OY	8125	tgcgttgatgtaacgactcttgggccacagcgcgtcttgtaataaataactcttgctttggc	8184
Db	8068	TGCGTGTGTACGACTGTGGGGCCCAAGCGGGCTTGGAAATAAATATCTTGTGCTTTTTC	8127
OY	8185	atcaa	8189
Db	8128	ATCAA	8132

RESULT	12
AX052638	
LOCUS	AX052638
DEFINITION	Sequence 7873 bp DNA PAT
ACCESSION	AX052638 from Patent WO0071726.
VERSION	AX052638.1 GI:12226828
KEYWORDS	
SOURCE	porcine endogenous retrovirus.
ORGANISM	
REMARKS	

1 (bases 1 to 7873)

methods to inhibit infectious agent transmission during
xenotransplantation
Patent: WO 0071726-A 23 30-NOV-2000;
Mayo Medical Ventures (US)

FEATURES	Location/Qualifiers
source	1. .7873

```

BASE COUNT      2200 a      1914 c      1961 g      1798 t
ORIGIN           /organism="porcine endogenous retrovirus"
                  /db_xref="taxon:61673"

```

Query Match	68.7%;	Score 5638.8;	DB 6;	Length 7873;
Best Local Similarity	84.8%;	Pred. No. 0;		
Matches 6718;	Conservative	0;	Mismatches 782;	Indels 420;
				Gaps 19;

[illegible]

Db	301	GGCCCTATCTTACGTGCGACGATTGTGGACAGAGATCTCCGGCATGGTTAACTTG	366
QY	887	gctcaataaaccaagaagaacgcagagctccccgaatctcgtgctcttggagagaaacaaca	946
Db	361	GCTGATTAATACCCAAAGAAAGCCAGGTCCTCCGAATTCCTTGCTTTGGAGAGAAAACAACA	420
QY	947	ctctggcgaataaagctcgaaccctctctctgctatctaaccccgagatcgaagccgcga	1000
Db	421	CTCGGTGAAAAAGTCAGACGCTC-TGCTCATATCTACCCGAGATTTGAGAGCCGCCG	479
QY	1007	cttggccggaaccccaacctgttcccacccccctatccagcaaggtgtcttggagg	1066
Db	480	CTTGGCCGGAACCCCAATCTGTCTCCCAACCCCTTTATCTGGCCACAGGCTGTGCAAGG	539
QY	1067	ga-ccctcgcgcctctctggagctccggtgtgtggaggagacctgcgcgcggaaatccgaagcc	1112
Db	540	GACCTCTGCCCTCTCTTGAGACTCCGCGGTGGAGGAACTCGCTGCAGGGAGACTCGAGCC	599
QY	1126	ggaagagcgcaacccccgagcgagcgaacagacgagatcgcgatattacgcctcgcacatg	1188
Db	600	GGAGGGGGCGCCACCCCGGAGGAGACGAGATCGGACATTACGCTGCGCAGCTACG	659
QY	1186	ggccctcccatcgcaggggggccaatttgcagccctcagatattggcccttctctgcag	1244
Db	660	GCCCTCCCAACACCGGGGGCCAAATTGACGCCCTCCACGATTTGGCCCTTTCTTCGCA	719
QY	1246	atctcataatltggaataactaacatccctctctcgagagatctcccaagcgtcagag	1300
Db	720	ATCTCTAATTTGGAGAAACTACATCCCTCTTCGAGAGATTCGCCAACGCTCACGG	779
QY	1306	ggttgtgtgagtcacctatgtctctcaacagctacttggagatgtgtcaacagctcgc	1366
Db	780	GGTTGGTGAGTCCCTTATGTCTCTCACGACCTACTTGGATGATTCACACAGCTGC	839
QY	1366	tgcacagacctctcaacaacgagagagagagaaattcgttaagagcttaagaaaaatgt	1422
Db	840	TGCAGACACTTTTACAAACGAGGAGGAGAGAGAAATTCGTTAGGCGCTTAAGAAAAAG	899
QY	1426	ttccctggcgacgagcgagcccaacgagatgtgcaaaatagatltgacagtggaattccct	1485
Db	900	TTTCCGGGGCGGACGGCGACCCACGCACTTGCAGAAATGATGTGACATGGGATTTCCCT	959
QY	1486	tgcattgccccggttggagactaaacaacgcttgaagtgtagggagagcttgaatactatc	1545
Db	960	TGACTCGCCCCGGTTGGGACTTACAAACGCGCTTAAAGTAAAGGAGACCTTGAATAACTATC	1019
QY	1546	gcacagctctctgtgtgcgagctctccgagcgccctcaagaagcgcccaactaatttggctaaag	1605
Db	1020	GCCAGGCTGTGTGGCGGGTCTCCGGGGCGCTCAAGAGAGGCCCACTAATTTGGCTAAAG	1079
QY	1606	taagaagagtgtatgcaggaacgaaacgaacactccctcgatattctcttggagagctcag	1665
Db	1080	TAAAGAGAGTGATGCAAGGAGCCAAATGAACCCCTCTGTTTCTTGAGAGGCTCTTGG	1139
QY	1666	aagcttcagagcgatcaaccccttttgatctactactccagagccccagaagcctcagtg	1725
Db	1140	AAGCTTCAGGCGGTACACCCCTTTTGTATCCACTCAAGGCCCCAAAAAGGCTTCAGTG	1199
QY	1726	ccctgcgctcatcttggcgagctgcgctcttgcgatatacgaagaagaactcagaagactgagag	1785
Db	1200	CTTTGGGCTTTATAGGCACTGACGCTTGATATTATTAAGAAAGAGCTTCAGAGACTGAGAG	1259
QY	1786	ggttaacaggggcttgagttactgtactagtgtagagagggcagagaagtgattatacaga	1845
Db	1260	GCTTACAGGAGCGGTGAGTTACGTATCGTAAAGAGGCGACAGCAAACTATTATACAAAA	1319
QY	1846	ggagagacaagagagagagaaacagagagaaagaaagagagagaaagaaagagagagaa	1905
Db	1320	GGGAGAGAGAGAAAGAAAGGACAAACAAAAGAGAGAGAAAGAGAGAAAGGGGAGAA	1379
QY	1906	gacgtgatagacggcaagaagaatttgcactaagaatcttggccgcagatgtgttaaagga	1965
Db	1380	GAGGTAAATTAACCGCAAGAGAAATTTGACTTAAGATCTCTGGCTGTCAGATGTTAAAGGA	1439

QY 1966 agagcagcaggagagagagagagatlttagaataatagtgcaagcccttagacagtcag 2025
| | | | |
Db 1440 AAGCAATATCGGAAAGAGAGAGATTTAGGAAATTTAGGTGAGGCCCTTAGACAGTCAG 1499
QY 2026 ggaacccgaggaatagagacccactcgacaagacagtggtggtlatgttaaaagaaag 2085
| | | | |
Db 1500 GGAACCTGGCAATAGAGCCCTCGACAGAGACCAATGTGCAATTTGTAAGAAAAAG 1559
QY 2086 gacactggcgaaggaactcgtccccaagaagagaaacaaagacgaaagtcctagctag 2145
| | | | |
Db 1560 GACACTGGGCAAGGAACTGCCCAAGAAAGGAAACAAAGAGACTGAAGGTTTACTCTGAG 1619
QY 2146 aagaagataaagattagggagagaggggttcgagacccctcccgagcccgagtaact 2205
| | | | |
Db 1620 AAGAAGATTAAGACTAGGAGAGAGCGGTTGAGCCCTCCGAGCCCGAGGTAACCTT 1679
QY 2206 tgaagtgagagggcaaccagttgagtcctggttgatataccgagcgagacatcagtc 2265
| | | | |
Db 1680 TGAAGGTGAGGAGGCAACCACTGAGTTCTGTTGATACGGGAGCAACATTCAGTGC 1739
QY 2266 tgcatacaacatlaggaaactaaagaaaaaactcctgggtgagtggtgcacagggc 2325
| | | | |
Db 1740 TACTACAGCATTTAGGAAACTAAAGATTAATAAAATCCTGGGTGATGGGTCCACAGGCG 1799
QY 2326 aacggcagatlatcagagatcccgaaagacggtgacttgagggtgaggggttaacc 2385
| | | | |
Db 1800 AACACAGATTCATGAGCTACCCGAGAGAACGTTGACTTGAGGATGGAGGGGTAAACC 1859
QY 2386 actcgttctgtcatccctgagtgccagtaaccctctagtagaagtaacttaactgaca 2445
| | | | |
Db 1860 ACTCGTTCTGTCATACCTGAGTCCGAGACCCCTCTTAGGTAGACATTTATGACCA 1919
QY 2446 agatggagagctcaaaatlcttcttgaaacaagagagacccagagtgctgtgatacaac 2505
| | | | |
Db 1920 AGATGGAGACCAAAATTTCTTTGAAACAAGGAAACCAAGAGTGTCTCAATAAACAAC 1979
QY 2506 ccatactgtgtgacccctcaatagatagatgaaatcagatatactcccaagtaa 2565
| | | | |
Db 1980 CTATCATCTGTGTGACCTCCCAATTAGATGACGAATATGCACTATACCTCCCTAGTAA 2039
QY 2566 agcctatcaagataatacagtcctgtgtgagcagttccccaagccttgagcagaacg 2625
| | | | |
Db 2040 AGCCTATCAAAATATACAAATCTGTGTGAAACAGTTTCCCAACCTGGGCAAGAACCG 2099
QY 2626 cagggatgggttgagcaagcaagttccccaaggttaataactgaaagccagtgcta 2685
| | | | |
Db 2100 CAGGGATGGGTTGGCAAGCAAGTCCGCCCAAGTTATTCACCTGAAGGCCAGTGCCA 2159
QY 2686 caccagatlatcagagagatccctctgagtagagagagcctgagaaagaaatctggcgc 2745
| | | | |
Db 2160 CACCACTGTGACGACAGCTACCTTCTAGTAAAGAGAGCTCAAGAGGAATTCGCGCCG 2219
QY 2746 atgtcaagatlaatacacaagggcatcagttcctgtgtccaatcccttgaaatact 2805
| | | | |
Db 2220 ATGTCCAAAGATTATTCACAAGGGCATCTTCTAGTTCTCTCAATCTCCCTGGAATATCT 2279
QY 2806 cccctgtacccgttagaagccttgagcaaatgatatcagaccagtaagagccttgagag 2865
| | | | |
Db 2280 CCTGTGACCGGTTAGAAAGCCCTGGGCAATGACTATGACAGCAGTACGAGCCTTGAGAG 2339
QY 2866 aggtcaataaagagtgagcagagacatacccaacggtcccgaaaccttataaccttla 2925
| | | | |
Db 2340 AGGTCAATTAAGCGGTGCGAGATATACCCCAAGACAGTCCGAAACCTTTAAACCTTTGT 2399
QY 2926 ggcgcctcccgctgaagcagagactgtlacaaglatltgacttaaaagatgccttctc 2985
| | | | |
Db 2400 GTGCTCTCCACCCCAACGAGAGCTGTATACAGTATTTGAAGGATGGCTTTTCTT 2459
QY 2986 gctctgagatatacccccactaagcaaacacttltgccttgaaatgagaaatccagta 3045
| | | | |
Db 2460 GCCTGAGATTACACCCCACTAGCAACCACTTTTGTGCTTCGATGAGAGATCCAGGTA 2519

QY 3046 cggagaagacccggcagctacacctgagaccgagctgcccaaggttcaagaaactcccca 3105
| | | | |
Db 2520 CGGGAAGAACCGGGGAGCTACCTGAGCCGACTGCCCAAGGGTTCAAGAACTCCCGCA 2579
QY 3106 ccacltltgaagagacccatacacaagagaccttgccaaacttcagatccaaacccctcag 3165
| | | | |
Db 2580 CCATCTTTGACGAAGCCCTTACACAGAGACTGGCCAACTTCAAGATCCAAACCTTCAAG 2639
QY 3166 tggacccctcccgatccgtgatgacttctcttgccgagagcccaacaaagactctc 3225
| | | | |
Db 2640 TGAACCTCTCCAGTACGTGATGACTGCTTTGTGGCGGAGGACCAACAAAGAGACTGCT 2699
QY 3226 tagaagtlacgaagcactactctggaatltgtcagactaagactaagcctctgcta 3285
| | | | |
Db 2700 TAGAAAGCAGAGAGGCACTACTGCTGAATTTGCTGACCTACAGAGCTCTGTGCTA 2759
QY 3286 agaagcccaagatttcagagagaggtlaacatacttgggtgatacattgggggggggc 3345
| | | | |
Db 2760 AGAAGGCCAGATTTGGAGAGAGAGGTAACTATCTTGGGGTACAGTTGCGGGAGCGGC 2819
QY 3346 agcgtgtgagcagagagcaggaagaaactgtatccagataccgcccccaacagag 3405
| | | | |
Db 2820 AGCGATGCTGACCGAGGACAGGAGAAACTGTAGTCAGATACCGGCCCAACCAACAG 2879
QY 3406 ccaacaagatgagagaglttggggagcagctgagatlttcagacgtgtgatacccggt 3465
| | | | |
Db 2880 CCAAAAGATGAGAGACTTTTGGGACACTGTGATTTTCAACACTGTGATCCCGGGT 2939
QY 3466 ttgcgacctagagagcccaactcaaccggttaacaaagaaagggggttctcctggg 3525
| | | | |
Db 2940 TTGGGACCTTAGACGCCCTTACCCACTTAACCAAGAAAGGGGAATTTCTCTGGG 2999
QY 3526 ctcttgagcaccagaaagagctltgatgtatacaaaagccctgtgagcagctgtc 3585
| | | | |
Db 3000 CTCTGTGAGCACCAAGAGCATTTGATGCTATCAAAAAGGCCCTGTAGGCGCACTGCTC 3059
QY 3586 tggccctccctgacgtactaaaccccttaactcttatagtgatgaggttaaggagtag 3645
| | | | |
Db 3060 TGCCCTCCCGAGGATTAACCTTTATCCCTTTATGTGAGATGAGCGTAAAGGAGTAG 3119
QY 3646 ccgagagagtttaaccacaacccctagagacatgagagagacgtgtgctactgtca 3705
| | | | |
Db 3120 CCGGGGAGTTTATACCCAAACTTAGAGACATGAGAGAGACTGTGCTACTGTGCA 3179
QY 3706 agaagctgtatccctgtagcagagtggtggccggtatctgaaagcctatcagcagctgtg 3765
| | | | |
Db 3180 AGAAGCTGATCTCTGTAGCCAGGTGTTGGCCCGATGCTGGAAGCTATGACGACTGTGG 3239
QY 3766 ccatactgtgtcaagagcgtgacaatgtacttgggaacagataataactgtatagccc 3825
| | | | |
Db 3240 CCATATCTGTCAAGAGCCCTGACCAAAATTTGACTTTGGGACAGAAATTAAGTAAATGCC 3299
QY 3826 cccatgtatgtgagaaactgttgcgagcccccagagccgtatgataccaacgcccga 3885
| | | | |
Db 3300 CCGATGCTGTGAGAGACTGCTTTGGGAGGCCCCAGACCATGTGATGACCAAGCCCGCA 3359
QY 3886 tgaaccaatlaaagcctgttctcaagagaggttcaacttgcctccacagcgcgc 3945
| | | | |
Db 3360 TGACCACTATCAAAAGCCTGCTTGCACAGAGAGGCTACGTGCTCCACAGCCCTC 3419
QY 3946 tcaacctgtcaactctctgcctgaagagactgtgaacccgtgacatgatgtgcatc 4005
| | | | |
Db 3420 TCAACCTCTGCACTCTTGTGCTGAAGAGCTGATGAACCAAGTATATATTATGCAATC 3479
QY 4006 aactatgatgtagagagactgggttccgcaagagacctatacagataccgctgtagag 4065
| | | | |
Db 3480 AACTATTATTTAGAGAGACTGGGTCCGCAAGACCTTACAGACATACCGCTGACTGAG 3539
QY 4066 aagtgtaacctgtgtcactgaagcagctatgtgtgtgaaaggttaagagagtgctg 4125
| | | | |
Db 3540 AAGTGTAACTGTTCACTGACGAGAACCACTATGTAGTGAAGGTAAAGGATGAGCTG 3599
QY 4126 gggcgagctgtgtgagcggagcccgacagatctgggcagcagctgtgcggaagaaact 4185

|||||
Db 3600 GGGCGGGTGTGGAGGAGCCCGACGATCTGGCCAGCAGCGCTGCCGGAAGAACTT 3659
OY 4186 cagcgcaaaagctgagctcatgcccacagcaagcttcgagctgcccgaaggaat 4245
Db 3660 CAGCACAAGAGGCTGAGCTCATGGCCCTCAGCAGAACTTTGGCGGTGGCCGAAGGGAAT 3719
OY 4246 ccataacattatagcagcagagtagtgccttgccagctgcacagctacacgagga 4305
Db 3720 CCATTAACATTTATACACAGCAGAGGTATGCTTGGCAGTGCACAGCTACACGGGCCA 3779
OY 4306 tctataaaccaaggggttgcttacctgaagcagggagggaaataagaacaagagaa 4355
Db 3780 TCTATAACCAAGGGGTTGCTTACCTTAGCAGGGAGGGAATTAAGAACAAAGAGAA 3839
OY 4366 ttctaaacctatagaagccttatacttgccaaaagagctatataacgtctctg 4425
Db 3840 TTCTTAAGCCTATTAGAACCTTTACATTGGCAAAAAGGCTTAGTATTATACACTGCTCTG 3899
OY 4426 gacatcagaaagccaaagatctcatatcagaagggaaccagatggtcagccggtgca 4485
Db 3900 GACATCAGAAAGCCAAAGATCCCATATCCAGAGGGAACAGATGGCTGACCGGGTTCGA 3959
OY 4486 agcagcgacccagagcttgtaaccttgcttgcctataatagaagcccaagccccagac 4545
Db 3960 AGCAGCGACGCCAGGCTGTAACTTCTGCTTATGATAGAACACCCAAAGCCCCAGAAC 4019
OY 4546 ccagacgagctacacccctagaagacgctgcagaagataaaagataagacagctctg 4605
Db 4020 CCGGACGACATACACCTTAGAAGACTGGCAAGAGATMAAAAGATGACAGTCTCTG 4079
OY 4606 agactcggagggagctgctatctaccctacatagtggaagaaatccctgccccaaag 4665
Db 4080 AGACTCCGGAAGGAGCTGCTATATCTCAATGGAAGAAATCTCCCAAGAAAG 4139
OY 4666 ggttagaatalgtccaaacagatacatgcttaaccacccaggaactaaacacactgagc 4725
Db 4140 GGTTAGATATGTCCACAGATACATGCTTAACCCACCTAGACACTAAACACTTGACG 4199
OY 4726 agttgttcagacatcccccatactgcttgaggtaccagagtggtcgtactggtg 4785
Db 4200 AGTTGGTCAGAAACATCTCTTATCATGTTCTGAGGGTACAGAGTGGCTGATTGGGTG 4259
OY 4786 tcaaacattgtgccccctgcagctgtgtaagtctaactccctccagaatacccccagaa 4845
Db 4260 TCAAAACCTGTGTCCCTGCCAGCTG----- 4285
OY 4846 agagactaagggaagccacccagcgctcactggaagtgcactcactgaagtaagc 4905
Db 4286 -----GTAAGC 4293
OY 4906 cggctaaatagcgaacaaatatactatgtgttttttagacacctttcagaatggtag 4965
Db 4294 CGGCTAAATAGGAACAAATATCTATGTGTTTGTAGACACTTTTCAAGATGGGTAG 4353
OY 4966 aggtcttacctaaagaaagaaactcaacgctgtgctgaagaaataactgga-gaga 5024
Db 4354 AGGCTATCTACTATAGAAAGAACTTCAACCGTGTGCTTATGAATAATCTGGAGGGA 4413
OY 5025 attttccaaagattggaatcactaagtaagggctcagacaatgctcagcttcgt 5084
Db 4414 ATTTTTCAGATTTGGAATACCTTAGGTAAATAGGGTCAAGACAAATGTCCACCTTTCGT 4473
OY 5085 gcccaagttaagtaaggtgcacaaagatactggtggaattgattggaactgcatgtgca 5144
Db 4474 GCCCAAGTAAGTACGGAGCTGGCCAGATATTGGGATGATTGGAATCGATGTGCA 4533
OY 5145 tacagaccccaagctcagagacaggttagagagatgaatagaacattaaagacact 5204
Db 4534 TACAGACCCCAAGCTCAGAGACGGTAGAGAGATGAATAAGCAATTAAGACACCTT 4593
OY 5205 accaaattgacacagagagctgcattaaagttgaltgagctctcctccttctgctt 5264
|||||
Db 4594 ACTAATTGACCGGAGAGACTGGCGTTAATGATTGCAATACCTCTCCCTGCTTGTGCTT 4653
OY 5265 tttaaggttaggaacacccctggaacagttggctgacccctataatgtctacagg 5324
Db 4654 TTTAGGGTTAGGAACACCCCTGGACAGTTTGGCTACCCCTATTAATTAATCTACGGG 4713
OY 5325 ggaacccccctgttgagaatctgcttgcctgaagtaagtgatgtgtgtcttccag 5384
Db 4714 GGAACCCCCCATTTGGTAGAATTTGCTCCGTACATAGTGTACCTGTGCTTCCAG 4773
OY 5385 ccttgtctctagtcacagcgctcagagtgaggtagagcagcgctggaagcagc 5444
Db 4774 CTTTGTGCTCTAGGCTCAAGCAGTGAAGTGGTGAGAACACAGACGAGGAGCACTC 4833
OY 5445 cggagagcctac-----agagagagcttgcaagctcacaatcgcttccagttgagat 5501
Db 4834 CGGAGGCGCTACACAGAGAGAGAGACTTGCAGATCCACATCGTTTCCAAAGTGGAGAT 4893
OY 5502 tcaagctatgttagagcaccgctgcaggaacccctgagacccgctggaagagactat 5561
Db 4894 TCAGTCTAGCTTAGAGCCACCTGCGAGAAACCTCGACACTCGGTGAGAGGCCCTTAT 4953
OY 5562 ctgtaactctgacacacacacagcgctgaaagtcgaaggaatcccaactgataccat 5621
Db 4954 CAGGTACTTTTGACACACACACAGCGCTGTGAAGTCAAGGAATCTCCACTGGATTCAT 5013
OY 5622 gaatcccaagttlaagccgagccacccctccagctgggggtggaagccgaagaactag 5681
Db 5014 GCATCCACAGTTTAGCCCGCGCCACCTCCGATGGGGTGAAGGCCGAAGAACTGAA 5073
OY 5682 aatcccttlaagcttgcctccatcgctcgtgtcttactcaacaataactccccagc 5741
Db 5074 AATCCCTTAAGCTTGCCCTCCATCGCGTGTCTTACTGTCAATTAACCTCCATGAT 5133
OY 5742 cagltagaagccttatagaagagcttgaaecccatagaacttaactcactgct 5801
Db 5134 TAAATGTAAGCCCTTGTGGACGCGCAACCCATTAACCTTATCTCTCATCTGCTT 5193
OY 5802 gattatagccctgatacagggctgacagtgtaaatagaacactgaggtgtgtcctcagag 5861
Db 5194 ACTTACTGACTCGGAGACAGTATTAATTAATTAACGACTCAAGAGGGAGGCTCCCTGGG 5253
OY 5862 caactggtgagctgaaactgacatcttgcctccagatgtaaacccgctgtta-----a 5915
Db 5254 GACCTGTGGCTGCAATTAATGTCTGCTCGATGAGTAATCTCCGTGTCAATGACCA 5313
OY 5916 aagcacactcccaacccagctcagtaagtgatgaggtcactgctgcagag-----cacaga 5972
Db 5314 GGCACACCCCGATGTACTCCGTGTACGGGTTTACGTTGCCAGAGACCCCAAA 5373
OY 5973 gaaagagaatactgtgagggttctgaggaaactcctcgtgagagatgagagctgcac 6032
Db 5374 TAATGAGAAATATGTGGAATTCCTCAGAGATTTCTTTTGGACAGCAATGGAGCTGCTAAC 5433
OY 6033 ctcaagatgagagctggaatgagcgaatctctcctccaggaacccggtlaaaatctcct 6092
Db 5434 TTTATATATGGAATGGAATGGCCAGTCTCAGCAAGACAGATAAGTAAGTACTCTT 5493
OY 6093 tgcattcc----- 6103
Db 5494 TGTTAACAATCTTACCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5553
OY 6104 gcccggaagtacaaatagaataactataagaataagagctgctcccatcagact 6163
Db 5554 GCACAGCGGGGTACAAAAGATG-TACGAATTAAGCAATTAAGTGTCAATTCCTTAAACC 5612
OY 6164 tagattatcaagataagtttactg-aaaggaacaggaataatctcaaaagtgga 6221
Db 5613 TAGATTACTTAAATAAGTTTCACTGAAAAAGAAAAAAGAAATAATTAATTAATTAATTA 5672
OY 6222 taatgtatgagctggaatagttttat-----tatgtgcggggagcaggtcca 6277
Db 5673 TAAATGTATGCTTTGGGGAATAGTATGAGAGCTCTGGAGAGAAAGAGATCTG 5732

QY 6278 ctttaaccattccgcttagatagagacgagagacacccctgtgcaatgagaccg 6337
 Db 5733 TTCTGACTATTCGCCCTCAGATAGAACTAGATGGAACCTCGGTTGCTATAGACAA 5792
 QY 6338 ataaagactgctgacagagagcccgccctgagaccacgactgacgag 6397
 Db 5793 ATAGGCTTTGGCCGAACGACCTCAATCCACGA----- 5829
 QY 6398 cccaattaacctgcctgagccctgacataacacgcccctagcaacagtlaccatgag 6457
 Db 5830 -----ACAGAGGCCATCTCCTTACCCCTCTGATTAACAATACAACTCTGAT 5876
 QY 6458 tgaattccacaacagccctagaaaccccaagtgctccgtgttaagacagacagac 6517
 Db 5877 CAGTCCCACTGA-----GCTTACATCACTATTAAACAGGGGCCAAAC 5921
 QY 6518 tcttcagctcatalccagagagcttccaaagcatalcaactccagaccctgacact 6577
 Db 5922 TTTTAACTCATCCAGAGAGCTTTTCAACCTCTTAACCTCCAGACTCCAGAGGCTACCT 5981
 QY 6578 ctctctgtgcttctctctcctcagagccctctattatgagggatgagctaaagaa 6637
 Db 5982 CTCTCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 6041
 QY 6638 gaaattcaatgtgacaaagacatagaaatcaatgatacagtgaggggtccgaataagc 6697
 Db 6042 GGAATTCATGTGACAAAGAACATAGAGCCATGTACATGGGATCCCAAAATATAGC 6101
 QY 6698 ttaacctactgaagttcccgagagagacatgacatagaaagagctcccccacc 6757
 Db 6102 TTACTCTTACTGAGGTTTGTGAAAGGCACTGTAGAGATGTTTCCCTCATCCACC 6161
 QY 6758 aacacctgtctatagtagctgtgtatagagcagccctcagaataatgattatgac 6817
 Db 6162 AACACCTTTTATGACACACAGAACCTTTATATCAACCTCTGAGAGTCACTATCTGTAC 6221
 QY 6818 ctgttataacagtggtgagcagatgacatgatacagtgatacccccctgtgttccacctg 6877
 Db 6222 CTGCTTATGACAGGTGCTGAGCATGTATATCTGATTAACCCCTTGTGTTTCCACCTTG 6281
 QY 6878 tcttcaaccaatccaagaattctgtgtcagtgatcccaatcgtcccgagtgatcc 6937
 Db 6282 TTTTCAACCAAACTAAAGACTTTTGGCTTATGCTTATGCTTATGCTTATGCTTATGCT 6341
 QY 6938 atcctgaggaagtggttctgtatgatactatgctatgatacagcccaaaagagac 6997
 Db 6342 ATCCGAAAAGAGCTCTTGTATGATATGATATGATATGATATGATATGATATGATATG 6401
 QY 6998 ccgatacccttaacctagctgataatgctgataagagcagccgtgtgagtagagacag 7057
 Db 6402 CCAATATCCCTGACACTGACTGATATGCTCGGATTTGGAGTGGCTGCGAGGGGTGGAAACAG 6461
 QY 7058 ggaacagctgcctgatalccagagacacagcagctagagaaagagactggtgagctacatg 7117
 Db 6462 GAAAGGCTGCGCTTATATCAAGAGACCGCAACAGCTGAGAAAGACTTATGATTAACCTATC 6521
 QY 7118 cggacatgacagaagatctccgagccttaagagatgctgttagcaactagaagagctcc 7177
 Db 6522 GAATTTGACGGAAGATCTCCAAACCTTAAAGAAATCTGTACATCACTGAGAGAAATCCC 6581
 QY 7178 tgaattctgtcgaagtgttctacagacagcgagggatlatgactctgttctaa 7237
 Db 6582 TAACCTCTTATCTGAAGTGTGTCTACAGAAAGAGAGGGGTTAGATCTTTATTTCTTA 6641
 QY 7238 gagaagtggttatagtgagccttaaaagaagaatgtgtctctatgtagactcag 7297
 Db 6642 AAGAAGAGAGATTTATGTGTGCTTTGAAGAGAGATGCTTTTATGTGATCATTTACG 6701
 QY 7298 gagcatcagagactcatalcaacagagcttagaaaaa----- 7334
 Db 6702 GGGCCATCAGAGACTCCATGACAAAGCTTATGAGAAAGACTGAGAGCCGCCCGGGTCT 6761

QY 7335 ----- 7334
 Db 6762 CCAACACCCAGACAGTGTGCTTGTTCATTTAAAGAACTGCAATGTAACCGCTGAGCT 6821
 QY 7335 -----gttagagagcgttcgaaggaagagagctgacagaggt 7375
 Db 6822 AAAGCCAGCTTAGAGAAAGGTTGAGAGCGGTGAGAGCGGCAAGGAAAGAAACTCAAGAGT 6881
 QY 7376 ggttgaagatggttccacaaggtctctgtatbaccacccctgtctctcagag 7435
 Db 6882 GGTTCAGAGGATGTTCAACAGTCTCTGTGTTGGTACCTACTTCTGTGTTAAACAG 6941
 QY 7436 ggcctctagtagtctcgtcctgtactaacagttgggctgtgtaataagttg 7495
 Db 6942 GACCTTATATGCT 7001
 QY 7496 tgccttgttgaagaagagatgagtgagtgacataatgtagttagttaggaacaglac 7555
 Db 7002 TTGCTTCAATTAAGAACCAATTAAGTCAAGTCAATGATGATGATGATGATGATGATGAT 7061
 QY 7556 aagcctctgagcccaag 7615
 Db 7062 AAAGCCCGCTTACG---AGAGAGAGCTGGCGCTGACTCTACCAAGTTTCAAGATTGAGACT 7118
 QY 7616 atlaacaagacagaagtgaggaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatg 7675
 Db 7119 ATTAAACAGAGAAAGTGGCGAATGAAAGATGAAAGATGAAAGATGAAAGATGAAAGATGAA 7178
 QY 7676 cgaagaatlaataaagcttaaatgcccccgagaaatcagagagagagagagagagagag 7735
 Db 7179 CAGAGAGATTAAGAAAGACCTTAAATGCTGAAATGCTGAAATGCTGAAATGCTGAAATGCT 7232
 QY 7736 aatagtagaagtgacacccctcctatgtgtcagagagagagagagagagagagagagag 7795
 Db 7233 --ATAGTAAAGATCATACTTTTGTGCTTTTAAAGGCTTGTCTGCTGCTGCTGCTGCTGCT 7290
 QY 7796 atlaacaagaatgagatga----- 7814
 Db 7291 CTGTTGAGAAAGGAGAAACAGCCCTGATGATGCTGATGCTGATGCTGATGCTGATGCTGAT 7350
 QY 7815 -----ctatgctatctgagatctcgttaactgacg----- 7848
 Db 7351 AACGCTCTTAACCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 7410
 QY 7849 -----gacatagaagaatgtatcacatgacagccctagtgac 7891
 Db 7411 GAAGTCAATTTGCTTAAGGAGACCCAGTAAGATGAGGATGACCAAAATGTTGAACACA 7470
 QY 7892 tatctcaactgcaatctgactctgcccagagacccagagatgagagagagagagagagag 7951
 Db 7471 TATCTTGTGACACACATGCTCTCCACCCAGCAACATGCGCAATGTTAATCTTAATA 7530
 QY 7952 tatcttaaaatgagtggtccag 8011
 Db 7531 CAATTTTAATTAATTTGTTGCTGACAGAGCGGCTCTGCAAGTTTAAATGCTGCT--- 7587
 QY 8012 tggagcgcgcctcctcagatattttaaagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 8071
 Db 7588 -----TTGTGATATTTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 7635
 QY 8072 aaccccaataaagctgtccgaatcccgaccccgagagagagagagagagagagagagagagag 8131
 Db 7636 AACCCCATTAAGAGCTGCTCCGAGCTCACACCTGGGGCGGCACTCTCAACCCCTCGTGG 7695
 QY 8132 tgaagacgtgtggcccgag 8191
 Db 7696 TGACAGACTGTGGGCCCGAGCGGCTGGAATTAATATCTCTGTGCTGTTGATCAACAGA 7755
 RESULT 13
 AX052634
 LOCUS AX052634 7362 bp DNA
 DEFINITION Sequence 19 from Patent WO0071726. 12-JAN-2001

ACCESSION	AX052634
VERSION	AX052634.1
KEYWORDS	GI:12226824
SOURCE	
ORGANISM	porcine endogenous retrovirus, porcine endogenous retrovirus Viruses; Retroid Viruses; Retroviridae; Mammalian type C retroviruses; 1 Mammalian type C virus group.
REFERENCE	1 (bases 1 to 7362) Federspiel, M.J.
AUTHORS	Methods to inhibit infectious agent transmission during xenotransplantation Patent: WO 0071726-A 19 30-NOV-2000; Mayo Medical Ventures (US)
JOURNAL	location/Qualifiers
FEATURES	1..7362 /organism="porcine endogenous retrovirus" /db_xref="taxon:61673"
BASE COUNT	1997 a 1821 c 1881 g 1663 t
ORIGIN	

Query Match	67.3%,	Score 5528.2;	DB 6;	Length 7362;
Best Local Similarity	90.5%,	Pred. No. 0;		
Matches 6117;	Conservative	0;	Mismatches 508;	Indels 134;
				Gaps 15;

[illegible]

QY	717	caacatctgaagcttggagagccacacagaggagaccttaattctgnaattacocggct	776
Db	1377	CGACACATTCGATCTTGATTTGGATGGCCATCAGAGGGAGCTTTAAATCTTGAGATTATCTGGCT	143
QY	777	gttaagagcaatcaatttctaagacttggaccgcgcctctcaatccctgaatacgaagccataatc	836
Db	1437	GTTAAAGSCAATTATTTTTCAGACCTGGACCCGGCTCTCATCTCCCATTCAGAGAGCCCTATATC	149
QY	837	cttaagcttgcaagaattttggcagaagaatccctccgcatauggtttaaacataygctaaataa	896
Db	1497	CTTACGTGGCAAGATTTTGGCAGAGGATCTCCGGCATGGGTAAACCTTGCGTAATAAG	155
QY	897	ccaagaagaacagagttcccgaatcccttgctctttttggagagaaaaaacacacttgcccgaa	956
Db	1557	CCAGAGAAACCGACAGTCCCGGATTTCTGGCTCTTGGAGAGAAAAACAACCTGGCTGTA	161
QY	957	aaagtcagagccctctctctglatatcccccagagactgagagccgcagcttggccgga	1011
Db	1617	AAAGTCAAGCCCTC-TCCATATATACCCGAGATTGAGAGCGCGCGGCTTGCGCGGA	1677
QY	1017	accccaacctgttcccccaacccccctatcgaacagagtgctgtgtgaaggga-cctctgcg	1077
Db	1676	ACCCCAATCTGTTCCCCCAACCCCTTATCTGGCAACAGGGGTGCGAGGGAGACCTCTCG	1733
QY	1076	ccctctctgagactccgctgtgtgtggagagacccgtgcgcggagactcgaagcccgagagcgc	1133
Db	1736	CCCTCTCTGAGACTCCGGCGGTGGAGGGACCTGCTGCAAGGACATCGAGCGCGGAGGGGCGC	1793
QY	1136	caaccgcgagcggaacagacagagatcgagataltacacgtctgcagactatgcccctccat	1195
Db	1796	CACCCGGAGCGGACAGACGAGATTCGGACATTACCGCTCGGACGTAAGGGCCCTCCAC	1855
QY	1196	gccagaggggccaattgcaagccccctccaglatatggcccttctctctgcagactcctataa	1255
Db	1856	ACCGGGGGGCCAATTGGACGCCCTTCACGATTTGGCCCTTTCTTCTGCAAGATCTATPAA	1915
QY	1256	tggagaaactaacaatccccccttctcggagagatcccccaagccttaagaggttttgga	1315
Db	1916	TTGGAAMACTAACCTATCCCTCTTCTGAGAGATCCCAACAGCCTTACAGGGGTGGATGA	1975
QY	1316	gtcccttatgtctctcaccagcactatgggatagtatgtgtcaacagctgtctcagaacat	1375
Db	1976	GTCCCTTATGTCTCTCACAGCCTACTTGGGATGTGTCAACACCTCTCTCAGACACT	2033
QY	1376	cttcaacaaccgagagacgagagagaattcttgatlaagagctagaanaaatgttccctgggc	1435
Db	2036	CTTCAACAACGAGAGAGCGAGAGAAATTTGTTAGAGGCTAGAAAAAATGTTCTTGGGGC	2093
QY	1436	ggaagggagaccacacgaagttgcaaaatggagatltgacatagggatttcccttgactcgcc	1495
Db	2096	CGACGGGACCCACGACGATTTGGAAAAATGATTTGACATGAGGATTTCCCTTACATCGGCC	2155
QY	1496	cggttggagactacaacacgcgcgtgaaggtlaagagagcttgaaaatcatatcgcacaagctct	1555
Db	2156	CGGTTGGGACTACAAACAACGGCTGAAGGTAGGAGAGACTTGAANAATCTATCCGCAAGCTCT	2215
QY	1556	ggttgcggtgtctccggggcgactcctaaagacggcccacaaatttgcttaagtlaagagaagt	1615
Db	2216	GGTGGCGGCTCTCCGGGGCCCTCAAGACGGCCCACTAATTTGGCTTAAGGTAAAGAGAGT	2275
QY	1616	gatcagagagaccgaagaacactccctccggtattctcttggaagctccaatgaaagccttaag	1675
Db	2276	GATCGAGGACCGGATACACCCCTCTGTCTTCTTGAGAGGCTCTTGGAAGCCCTTACG	2335
QY	1676	ggagltcaacccttltgatctactcactcagaagcccaagaagcctagtgccctggcctt	1735
Db	2336	GGGAGTACACCCCTTTTATCTCCACGTCAGAGGGCCAAAAAGCGCTCAAGTGGCTTTGGCTT	2395
QY	1736	cattgggcagtgcgctctctgataatacgaagaanaacttaagagacttgaagaggttacaaga	1795
Db	2396	TATGGGACAGCTCAGCCTTGGATATTGAGAAAGAAAGCTTCAGAGACTGGAAGGGTTACAGA	2455
QY	1796	ggtcaggttaccgatactgtgaagagagcgcaagaagtgatcttcacgaagaaggaacaga	1855

QY	4016	tggagagactggtggtccgcaagaccttcaagacataccgcgtgactggagaagtgtcaac	4073
Db	4676	tggagagactggtgggtccgaaggaccttaccagacatccctctgctggaagtgtttt	4735
QY	4076	ctggttccatgacggaaagcaactatgtgtgtgaagtgaaaggaatgtgctggggcggagt	4133
Db	4736	ctgggttccactgacggaaacacactatgtgtgtgaaggttaagagatgtgcttggggcgccgt	4793
QY	4136	ggtgtgacggagcccgacacgatctgtggccgaagccttgcggaaaggaacttcagcgcataa	4195
Db	4796	gggtgacggagccacccacagatcttggccacagacgccttccgaaggaacttcagcgcataa	4855
QY	4196	ggctgagctcatgtggccctccaagaagtttgcgtgtgcccgaagggaaatccataaacat	4255
Db	4856	ggctgagctcatgtggccctccaagaagtttgcgtgtgcccgaagggaaatccataaacat	4915
QY	4256	ttatacggacagcaagtgtacgttcttgcactgtcacacgttacaagcggggccatctataaca	4315
Db	4916	ttatccacagacagcagatgtgcttggcactgtgcacacgttacaagcggggccatctataaca	4975
QY	4316	aaggggtgtgcttacctcaagcaggagggaaataaagaacaaagagaaatcttaagcct	4375
Db	4976	aaggggtgtgcttacctcaagcaggagggaaataaagaacaaagagaaatcttaagcct	5035
QY	4376	attaaagcccttaacatttgcacaaaaagcgtgatatataacgttgcctgagacataaa	4435
Db	5036	atttaaacctttaatatttggccaaaagagcgttatattatpacctgtctcgacatcacaa	5095
QY	4436	agccaagaatcatalctatgaagggaaaccaagatgtgcacgggtgtgccaagcaagcagc	4495
Db	5096	agccaaagatcccatatttccaaaggaagcagatggcgtgacccgggttgccaaacagcagc	5155
QY	4496	ccaggtctttaaacttctgtcctataatagaagcgcacaaagcccaagcccaagaccagaagca	4555
Db	5156	ccaggtctttaaacttctgtcctataatagaagcgcacaaagcccaagcccaagaccagaagca	5215
QY	4556	gtacacccctagaagaactgtgcagaagataaaagaatagacagagttctcttgagacatccgga	4615
Db	5216	gtacacccctagaagaactgtgcagaagataaaagaatagacagagttctcttgagacatccgga	5275
QY	4616	ggggagactgtctatacctcataatgaggaaagaaatcctgtcccaacaaagaaggttagaata	4675
Db	5276	agggagactgtctatcctcacaatgaggaaagaaatcctgtcccaacaaagaaggttagaata	5335
QY	4676	tgccaacagatatacgtctctaaaccaactlagaactaaagccctgcagcaggttggctcag	4735
Db	5336	tgccaacagatatacgtctctaaaccaactlagaactaaagccctgcagcaggttggctcag	5395
QY	4736	aaacatccctatcaatgtctctgaggtctacagagtggtgcgtgactcgtgtgtccaacaattg	4795
Db	5396	aaacatcctcttattatcattgtctctgaggtctacagagtggtgcgtgactcgtgtgtccaac	5455
QY	4796	tggtccctgcgcagctggtttaaigtctaatccttccagaatlaacctccaggaagaagactaag	4855
Db	5456	tggtccctgcgcagctggtttaaigtctaatccttccagaatlaacctccaggaagaagactaag	5515
QY	4856	gggaagccaccacagcgctcaactgtggaagtgtaactcaactaggtlaaagccgggttaata	4915
Db	5516	gggaagccaccacagcgctcaactgtggaagtgtaactcaactaggtlaaagccgggttaata	5575
QY	4916	cgggaacaaatctatgtgttcttgtlagaacacttccagatlyggtlagaagcttacc	4975
Db	5576	cgggaacaaatctatgtgttcttgtlagaacacttccagatlyggtlagaagcttacc	5635
QY	4976	tactaaagaagaactcaaacgtgtgtgtctaaagaataactatggaagaatattttccaa	5035
Db	5636	tactaaagaagaactcaaacgtgtgtgtgtctaaagaataactatggaagaatattttccaa	5695
QY	5036	atttggatacctcaaggtatagaaggtlcaagacatgtgtccagcttctgttcccaagtaag	5095
Db	5696	atttggatacctcaaggtatagaaggtlcaagacatgtgtccagcttctgttcccaagtaag	5755

QY	5096	ttggagacctcggccaagaatcttgggaattggaatttggaagaactgcatgtgcatacagacccca	515
Db	5756	TCAGGAGCTGGCCAAAGATATTGGGGATTGATTGGAAATCGGATTTGTCTCATACACACCCCA	581
QY	5156	aagctcaagcagcagtlagagaagatgaatagaaccactlaaagaagacccttaccgaattgac	521
Db	5816	AAGCTCAGGACACAGCTAGAGAGGATGAATAGAACCTTTAAAGACCTTACTAAATTGAC	587
QY	5216	cacagaactcgtgcatataatgatttgaatggctctccctggcttggcttttaaglytga	527
Db	5876	CGGGAGACTGGGGTTAAATGATTGGATAGCTCTCCCTGGCTTTTACTTTTAAAGGTTAG	593
QY	5276	gaacaccccttgaaagatttgggcttbaacccctatagaattgtctcaaggggaaccccccc	533
Db	5936	GAACACCCCTGGACAGATTGGGGCTGACCCCTATGAATTAATCTTACAGGGGAGACCCCCCC	599
QY	5336	gttggcagaatattgccttctgcacagctgcagatgtcgtctgtcttccagcccttattctc	539
Db	5996	ATTGGTAGAAATTTGGCTTCGGTACATAGTGTGCTGAGCTGTCTTTCCAGCCTTTGTTCTC	605
QY	5396	taagctcaagcgcgtcgaatgtggttgaagcagcgagcttgaagaagctctccggaaagccta	545
Db	6056	TAGGCTCAAGGCACCTTGATGATGGGTGAGACACAGACGCGTGAGGCACTCCGGAGGCGTA	611
QY	5456	ctc---aggaagagactctgcaagtctcaactcgaatccgcttccaagtttggagattcagctagt	551
Db	6116	CTCAGAGAGAGAGGAGACTTGGACATTTGGCATTCACATCTGTTTCCAAGTGGGAGATTCACTTACGT	617
QY	5513	tagacgccaccctgtaggaagaacctcgagaactcggctggaaggaacctatctgtactttt	557
Db	6176	TAGACGCCACCTGTGAGGAACCTTCGAGACTCGGTGGAAGGGCCCTTATCAGCTACTTTT	623
QY	5573	gaaccaaccaagcgtctgtgaagctcgaagaatcccaaccttgatccatgcatcccaagct	563
Db	6236	GACCAACACCAAGCGGTGTGAAGAATCGAAGGATCTCCACCTGGATCCATGACATCCACGT	629
QY	5633	taagccggcgacactctccgcgaactcggggttggagagccgaagaagctgaatcccttaa	569
Db	6296	TAAAGCGGCGGCACCTCCCGATTGTGGGGTGTGAAGGCCGAAMAACATGAAATATCCCTTAA	635
QY	5693	gcttcgcctccaatcgccctggtctcttactctaaacaataactcccaagcgagtagaanaac	575
Db	6356	GCTTGCCTCCATCGCGTGGTTCCTTACTCTGTGTAATACTCTCTCAAGTTAATGTAAGAAC	641
QY	5753	gcctatagagaagctcgaaaccccatagaaacttatacccttaccctgcgtacattatgac	581
Db	6416	GCGTGTGAGACAGCCCGCAATCCCATTAACCCATTATCTCAGCTGGTTACTTACTAGCT	647
QY	5813	ctgtaagcgggtgctactcgttaaataagaactcgaaggtgttctctctagaagcaccctgtagc	587
Db	6476	CGGTAACAGGATTAATAATTAAACAGCACTCAAGGGAGAGCTCCCTGGGACCTGATGGGC	653
QY	5873	ctgaactcatttctgctctcgatattgtatbaacccgcgttga-----aaagcaaacctc	592
Db	6536	CTGAATTTATNGTCTGCTTCCTTGATGACATAATCCCTGTCTCAATGACCGGCCACACCCC	659
QY	5927	ccaactcagctcgtcagttatggttctatctatgctgcaccga-----cacagaagaagaaat	598
Db	6596	CCGATGTACTCCGHTGCTTACGGGGTTTACGTTTGGCCAGAGACCCCAATATATGAAGAT	665
QY	5984	actgtgggggttctcggggaaactcctctctagagaatgtagcgtcgtcaccctcaacgagt	603
Db	6656	ATTGGGAAATTCCTCAGAGATTTCTTTTGGCAAGCAATGAGAGCTGCGTAATCTTATATGATG	671
QY	6044	gagacttgaaatgtagcattctctccagagaccgggttaaatcttcccttgt-----	609
Db	6716	GGAATGGAAATGGCCAGTCTCTCCACAGACAGAGATAGTTACTTTTGTTAACAATCT	677
QY	6096	-----caatccggcccgggcgaag	614
Db	6776	CTACACGTTAATATCAATTTAATTATGGCCATGGGAGATGGAAGAAATTTGGCAACACGCGG	683
QY	6115	tacaanaatgtaaacatataataagaatgaagcgtgcctcccatcgaactttagattatata	617

QY 361 ccaacctggggaagcggccggggaagtggggaagcgaagcgctggtgctctact 420
|||||
Db 371 CCACCTTGGGGAGCGCCCGGAGGTGGGAGAGCCAGGAGCGCTGGTGGTCTCTACT 430
QY 421 gtcggtcaagaagcagadttctgtgttgaaagcgaagctccctccggcgccgtccg 480
|||||
Db 431 GTCGGGTCAAGAGACCGAGTCTGTGTGTAAGCGAAGCTCCCTCCGCGCGCTCCG 490
QY 481 actctttgctgctctgtggaagcagcgagcggtcggtgtgtctgtatctgtgtt 540
|||||
Db 491 ACTCTTTGGCCCTGTTGTGGAAGAGCGGAGCGGCTGGCTGTGCTGTGATCTGTGTT 550
QY 541 ctgtctgtgtgtctctgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 600
|||||
Db 551 CTGTTCTGTGTGTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 610
QY 601 tgactaccccttgatttgactctgcagcacttgagcgaagttagatccagggctata 660
|||||
Db 611 TGACTACCCCTTAGTTGACTCTGCAGCCATTGGACTGAAGTTAGATCCAGGCTCTATA 670
QY 661 attgtcagttcaagtttaagaaggaccttggcagacttctgtcctctgaatgycgaa 720
|||||
Db 671 ATTGTCTAGTTCAAGTTAAGAGGAGACTTGGCAGACTTCTGTGCTGTGAATGGCGAA 730
QY 721 ccttcgatgtgtgatacagcacaagaaggacacttaattctgaatatctgtgtt 780
|||||
Db 731 CATTCGATGTTGATGAGGCGCATCAGAGGAGCCTTTAATTCTGAATTAATCGGCTCTTA 790
QY 781 aggcacatcatcttcaagctgacccggtctcactcctgactcaagagccctatactta 840
|||||
Db 791 AGGCATCATCTTTCAGACTGAGCCGCGCTCTCATCTGATCAGAGCCCTATATCTTA 850
QY 841 cgtggaagatcttgccagaagaacctccgcacatggtttaaaccatggctaataacaa 900
|||||
Db 851 CGTGGCAAGATTTGGCAGAGAGATCTCGCCATGGTTAAACCATGCTTAATTAACCAA 910
QY 901 gaaagcaggtcccccgaactctgtctcttggaagaaacaaacacatctggccgaaag 960
|||||
Db 911 GAAAGCCAGGTCCCGAATCTGTGCTGTGAGAGAAACAAACACTCGCGCGAAAG 970
QY 961 tcgagccctctctctglatatcaccggagatcgagagcgccgacttggccggaacc 1020
|||||
Db 971 TCGAGCCCTC-TCTCGATCTACCCGAGATCGAGAGCGCCGACTTGGCCGAAACC 1029
QY 1021 caacctttcccccaccccttalcagcagaaggtgtgtgtgagggga-cctctgcccc 1079
|||||
Db 1030 CAACCTGTTCCCGACCCCTTATCCAGCAGAGGTGTGTGAGGGAGCCCTCTGCCCT 1089
QY 1080 cctgagctcctggtgtgtgagggagcctgtgtcgggagactcgagcgagagagcgacac 1139
|||||
Db 1090 CTTGAGACTCTCGGTGTGTGAGAGACTCTGTGCGGAGCTGGAGCCGAGAGGCGCCACC 1149
QY 1140 cggagcaggaagaagagatcgagatactacgctgcgcacactatggccctccatgcca 1199
|||||
Db 1150 CCGAGCGGAGAGAGAGATCGGATATACCGCTGCGACCTATGSCCTCTCCATGCCA 1209
QY 1200 gggggccaattgagcccccgcglatatggccctttctctgcgagatctataattgg 1259
|||||
Db 1210 GGGGGCCAATTGAGCCCTCTCAGATATGTCCTTTCTCTGCAATCTCTAATATGG 1269
QY 1260 aaaactaacatcccccttctcggagatcccccaagcctcaggggtgtgtgtgttc 1319
|||||
Db 1270 AAAACTAACATCCCTCTTCTCGAGAGATCCCAACGCTCTCAGGGGTGGAGATCC 1329
QY 1320 ctatgttctctcaacagcacttggatgtatgtcaacagctgtcagacactctc 1379
|||||
Db 1330 CTATGTCTCTCAACGAGCTACTGTGAGATGATTGCAACAGCTGTGAGAGACTCTTC 1389
QY 1380 acaacggagagcagagagagatctgtttagaggtcagaaataatgttccctgggagcag 1439
|||||
Db 1390 ACAACGAGGAGGAGAGAAATTCGTTAAGAGCTTAAAAAAAATGTTCTGGGCGCGAG 1449
QY 1440 gggcgagccagcaggttgcaaaatgagatggaatgttcccttgctcgtcccggt 1499

Db 1450 GGCGACCCACGAGTTCGAAATGAGATTGACATGGGATTTCCCTGACTCCCGCGGT 1509
QY 1500 tgggtactaacaacagcgtctgaaggttagggagagcttgaatactatcgccagctgtgt 1559
|||||
Db 1510 TGGGACTTAACAACGGGTGAAGGTAGGAGAGCTTGAATAATCTATGTCAGCTCTGTG 1569
QY 1560 ggggtctccggggcgctcctaagaagcccaacttttgctgaagttaagagaggtgagt 1619
|||||
Db 1570 GCGGCTCTCCGGGGCGCTTAAGAGCGCCACTTAATTTGGCTAAGGTAAAGAGGTGATG 1629
QY 1620 caaggacgaagcgaacccctcgtglatctctggaagagctcaatggaagcctcagggcg 1679
|||||
Db 1630 CAGGAGACGAAGCAACCTCCCTCGTATTTCTTGAGAGGCTCATGGAACCTTCAGCGG 1689
QY 1680 ttaccccttttgatctactactaagagggcccaagaagcctcagttgctgtcctcalt 1739
|||||
Db 1690 TTCACTCCCTTTGATCTCTACCTTGAGGCGCCAGAAAGCCCTCAGTGCCCTTCATT 1749
QY 1740 gggcagtcggtctgtgatatcaagaagaacttcaagaagcttgaaggtttacagaaggt 1799
|||||
Db 1750 GGGCAGTCGGCTCTGTGATATCAGAGAAACTTCAAGACTGGAAGGTTTACAGAGGCT 1809
QY 1800 gagttacgtatctatgtgagagagcagaaggtgttatacagaaggagacagaagag 1859
|||||
Db 1810 GAGTTACGTGATCTAGTGAAGAGCGAGAGAGGTGTATTACAGAAAGGAGACAGAGAG 1869
QY 1860 gagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 1919
|||||
Db 1870 GAGAAGGAACAGAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1929
QY 1920 caagaagaagaatttgactaagatctgtgcgcagtggtgtgaaggagagagcagcagggag 1979
|||||
Db 1930 CAAGAGAAACAATTTGACTTAAGATCTTGGCGCAGTGTGTTGAAGGGAAGAGCGAGGAG 1989
QY 1980 agagagagagatcttagaanaatltagttcagccctagaagctcagggaaacctggcaat 2039
|||||
Db 1990 AGAGAGAGAGATTTTGAAGAAATTAAGTCTAGCCCTTACAGACTGAGGAACCTGGCAAT 2049
QY 2040 aggaagcccaactcgacaagaagcaggtgtgtgtgtgtgaagaagaagaagaagaagaaga 2099
|||||
Db 2050 AGACCCCACTGACAG 2109
QY 2100 aactgcccagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 2159
|||||
Db 2110 AACTGCCCAAGAT 2169
QY 2160 tagggagagagaggtgtcggagccctcccgagccagaggttaacttgaaggtgagggg 2219
|||||
Db 2170 TAGGGAGAGAGGGGTTCGAGACCCCTCCCGAGCCAGGTTAACTTTGAAGGTGAGGGG 2229
QY 2220 caaccagttgaattctcgtgtgatacggagagagagcaatcaggtgtctcaaacatata 2279
|||||
Db 2230 CAACCAAGTTGATTTCTGTTGATATACCGAGCGAGACTTCAAGTCTCTCAACCATTA 2289
QY 2280 ggaanaactaagaanaaanaaactcctgggtgtgtgtgtgtgtgtgtgtgtgtgtgt 2339
|||||
Db 2290 GGAANAATTAAGAAAGAAAGAAATCTGAGGTGATGAGTGCACAGGGCAACGGCAGATCCA 2349
QY 2340 tgactacccgaagaagcaggttgacttggaggttggagcgggttaacccatgttctgtgtc 2399
|||||
Db 2350 TGAGCTACCCGAGAGAACCGTTGACTTGGAGCTGGAGCGGTACCCACTTCTTCTGTGTC 2409
QY 2400 atccctgt 2459
|||||
Db 2410 ATCCCTGAGTGTGCGGAGTACCCCTTGTAGGTAGAGACTTACTGACCAAGATGGAGCTCA 2469
QY 2460 attcttttgaacaaaggaagcagaaggtgtctgtgtgaataacaaacccatcactgtgtg 2519
|||||
Db 2470 ATTCTTTTGAACAAGAGAGAGAGAGAGAGTGTGTAATTAACAACCATCATCAGTGTG 2529
QY 2520 accctcaatagatgagatggaatatcagactatattctcccaagtaagcctgataagat 2579
|||||

Db 2530 ACCCTCAATTAGATGATGAATATCGACTATATCTCTCCCAAGTAACCCGTGATCAAGAT 2589
OY 2580 ataccgctctgttggagcaagtttcccaagccttggcaagaacccgaaggatggttgg 2639
Db 2590 ATACAGTCTGTTGGACAGATTTCCCAAGCTGGGAGCAAAACCCGAGGATGGTTTG 2649
OY 2640 gcaaaagcaagttcccccacaggtatttaactgaagcgcaatgtttaacagttatagtc 2699
Db 2650 GCAAAAGCAAGTTCCCCACAGATTATTCACCTGAAGGCCAATGCTACCCAGTATCCAGTTC 2709
OY 2700 agacagtaacccttgaatagagagagctcgagaagaaatlttggccgcatgttcaagaalta 2759
Db 2710 ACACGATACCCCTTAGTAGAGAGAGCTCGAGAAGAAATTTGGCCGCAATGTTCAAAAGATTA 2769
OY 2760 atccacagagagcatctagttccgttccatcccttggaaatcccttccctgtaccggtt 2819
Db 2770 ATCCAAACAGGCGATCTCTGATTCCTTCCATTCCTTGGAAATACCTCCCTGCTACCCGCTT 2829
OY 2820 aggaagccttggagcaaatgattatcgaccagtaacagagcttgaagagaggtcaataaag 2879
Db 2830 AGGAAGCCTGGAGCAATGATATATGACCAAGTACAGAGCTTGAGAGAGGTCAATTAAGG 2889
OY 2880 gtgcagagacatacccccacggtcccgaaaccttaaacctcttgaagccttccgct 2939
Db 2890 GTGCAGAGATACACCAAGGTCGCCGAACCTTAAACCTTGAGCGCCCTCCGCT 2949
OY 2940 gaacggaacttggtaacagatttgaacttaaaagatgccttcttccctgaatlaaac 2999
Db 2950 GAACGGAATCGGTACACATGATTTGACTTAAAGATGCTTCTTGGCTGAGATTAC 3009
OY 3000 cccacacagcaaacacttlttgccttgaatggagagatccaggttgaaggaacgg 3059
Db 3010 CCACACTAGCAACCACTTTTACCTTCGATGAGAGATCCAGGTACCGGAAGAACCCGG 3069
OY 3060 cagctacacttggaccgcaacttgcacaaagggttcaagaactcccgcaacatcttgaagaa 3119
Db 3070 CAGCTACCTGGACCCGATGCCCCCAAGGTTCAAGAACTCCCGCACATCTTTGACGAA 3129
OY 3120 gccctacagaggaacttgcacaaacttcaagatccaaacacctcaagttgaccttccag 3179
Db 3130 GCCCTACACAGGAGCTGGCCAACTTCAGATCCAAACCCCTGAGTGACCTCTCCAG 3189
OY 3180 taagtgaatgaacttcttcttggcgaggaacacaaacagagacttcttaagaagtacgaag 3239
Db 3190 TAGGTGATGATCTGCTTCTGGGGGAGGACCCAAACAGAGATGCTTAAGGTACGAG 3249
OY 3240 gcaactcgttgaatttcttgaactgaagctacagaagccttcttgaagaaggccagatt 3299
Db 3250 GCACACTGCTGGAATTTGCTGACCTAGGCTACAGAGCTTGGCTAAGGAAGGCCAGATT 3309
OY 3300 tgcagagagaggttaacatacttgggttgaagtttgcggggcgaggaagagatgagctagc 3359
Db 3310 TGGAGGAGAGAGGTACATATCTTGGGGTACAGTTTCGGGGCGGAGGAGATGGCTGACG 3369
OY 3360 gaagcagagagaacttgaacttgcagataccgcccacaaacacgaacccaacagttgaga 3419
Db 3370 GAGGACAGGAAAGAACTGTAGTCCAGATACGGGCCCAACACACAGCAAAACAGTGAAG 3429
OY 3420 gagtlttgggaacgttgaatttgcagacttgaatccggggttgcgaccttaaga 3479
Db 3430 GAGTTTTGGGACAGCTGATTTTGGAGACTGTGATCCCGGGTTTGGGACCTTAGCA 3489
OY 3480 gccccactaccgcttaacaaagaaaggggagtttcccttgggtcctctgagcaaccag 3539
Db 3490 GCCCAGCTTACCCGCTTAACCAAGAA 3516
OY 3540 aagcatttgaatgatacaaaagggccttgaagcgaccttgccttgcctccctgac 3599
Db 3517 3516
OY 3600 gtaactaaaccttacccttattgtgtgagtgagagagaglagccgagagatttla 3659
Db 3517 3516

OY 3660 acccaaaccttagacatgagagagacctgttgcctacctgtlcaagaagattgalcct 3719
Db 3517 3516
OY 3720 gtagcagatgttgcgcccgtatgtcgaaggtatcgcagctgtgcctatactgttcaag 3779
Db 3517 3516
OY 3780 gaagctgacaattgacttctgggaagaataataactgttaatagcccccatgtcag 3839
Db 3517 3516
OY 3840 aacatgttcgcagccccacagacgaatgatatcaacaacgcccgcaltgaccatca 3899
Db 3517 3516
OY 3900 agcctgtcttcacagagaggttcaacttctgcctccacagcgccttcaacctgcact 3959
Db 3517 3516
OY 3960 ctctgcctgaagaagacttgaacacagtgatcaltgatgttccatcaactatgttag 4019
Db 3517 3516
OY 4020 gagactgggttcgcgaagacacttacagacataccgcttgacttggagaagtgttaacctg 4079
Db 3517 3516
OY 4080 ttcaactgaaggaagcagctatgtgtgtaaggttaagagatgtgttggcgagctgtg 4139
Db 3517 3516
OY 4140 gacggagccgcagatcttggccagcagccttgcggaagaagtaactlcaagcnaaagct 4199
Db 3517 3516
OY 4200 gagctatgcccctacgcgaagcttgcgcttgcggaaggaatacatalaacttatt 4259
Db 3517 3516
OY 4260 acggaacagatgatagtccttgcagctgcacagctacaggggcatctataaacaag 4319
Db 3517 3516
OY 4320 gggttgttacctcagcagaggaagaaataagaacaagaagaatcttaagctatla 4379
Db 3522 GGGTTGCTTACTTCAGAGGGAGGAAATTAAGAAACAAAGAGGAATTTCTAAGCTATT 3581
OY 4380 gaagccttaccttgcacaaagcctagctatatacaactgttctctgagacatcagaagcc 4439
Db 3582 GAAGCCTTACATTTGCCAAAAAGGCTACTATTATACCTGCTGGACATCAGAAACCC 3641
OY 4440 aaagatctcatatcttaagagggagacagatgtgttgcgggttgcacgaacgaagccag 4499
Db 3642 AAAGATCTCATATTCAGAGGAGACCATGCTGTGACCGGGTTGCCAAGACGACCCAG 3701
OY 4500 gctgttaaccttctgcctataataganaagcccaagccccaagacccaagacagctac 4559
Db 3702 GCTGTTAACCTTCTGCTATATAGAAAGGCCCAAGGCCAGAACCCAGACGATGAC 3761
OY 4560 accctagaagacttgcgaagagataaaagataagcgaagttcttgaacatccggaggg 4619
Db 3762 ACCCTAGAAAGACTGGCAAGATTAAGAAAGATAGACCAAGTTCTGAGACTCCGAGGGG 3821
OY 4620 acctgtataacctcatatgtgaaggaatcttgcgcccaaaaggaaggltgaatatgtc 4679
Db 3822 ACCTGCTTACCTCATATGGAAGGAATCTGCCCCAACAAAGAGGGTTGAATATGTC 3881
OY 4680 caacagatacatcgttcaacccaactagaactaaacacctgcagcagttgttgaaga 4739
Db 3882 CAACAGATACATGCTTAACCCACCTAGGAACCTAAACACCTGACAGTGGTGAAGACA 3941

Db 5971 GGGCATGTATCTGATTAACCCCTTGTCTTCCACCTTGGTTTTCACCAACCACTAAG 6030
 QY attctgtgcatgtgctcaaatcgtccccgagtgtaacacactcgtgaagtggtcc 6955
 Db 6031 ACTTTGGCTATGTGTCACCAATTTGCCCGGGGTGACTACTATCCGAAAGCAAGCTGC 6090
 QY ttgataatgactatcgatataacccgcaaaagagaacccgtatcccttaacctag 7015
 Db 6091 TTGATGATATGATGATGATATATCGGCCAAAGAGCCCATATCCCTGACACTAG 6150
 QY ctgtaatgctcgattgagcgccgctgagctgaggaacagagcgccctgtatca 7075
 Db 6151 CTGTAAGCTCGATGGAATGGAGTGGCTGACGGCGTGGCAACGAGACGGCTCCCTATCA 6210
 QY caggacacagcaagcaagagaagagacttggtagctacacatgcygccaatgacagaatc 7135
 Db 6211 CAGAGCCGCAACACTGTGAGAGAGAGACTTAGTAACCTACATCAATGTAAGGAGATC 6270
 QY tccgagccttaagagagctgtgttagcaacctagaagagtccttctgtctgaag 7195
 Db 6271 TCACAGCCCTAGAAAATCTGTGATGATGATGATGATGATGATGATGATGATGATG 6330
 QY tggcttctacagcaacgag 7255
 Db 6331 TGGTTCTACAGACAG 6390
 QY cagccttaagagaagatgtgtctctctctctctctctctctctctctctctctctct 7315
 Db 6391 TAGCCTTAAAG 6450
 QY tgaacaaacttagaagaaag 7375
 Db 6451 TGAGCAACCTTAG 6510
 QY ggttgaagagatgtgtct 7435
 Db 6511 GGTGGAAG 6570
 QY gggccctagtagctctgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgt 7495
 Db 6571 GGGCCCTAGTAGCTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 6630
 QY ttgcttctgt 7555
 Db 6631 TTGCTTGT 6690
 QY aagcctctgt 7615
 Db 6691 AAGGCTTGT 6750
 QY attacaagacaag 7675
 Db 6751 ATTAAACAAAG 6810
 QY ccaag 7735
 Db 6811 CCGAG 6870
 QY aataagtagaag 7795
 Db 6871 AATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6930
 QY ataacag 7855
 Db 6931 ATACAG 6990
 QY agaagaatgattatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcat 7915
 Db 6991 AGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7050
 QY tgcacag 7975
 Db 7051 TGCCAG 7110

QY 7976 agcgcgagctctcgatattttaaagatgtgtccatgagcgcgagctctcgatattt 8035
 Db 7111 AGCGGGGCTTCGATATTTTAAATGATTTGGTCCATGAGAGCGGGGCTCGATATTTT 7170
 QY 8036 aaatgattgt 8095
 Db 7171 AAATGATTTGTTGT 7230
 QY 8096 cgcagctgggagcgagct 8155
 Db 7231 CCGCACTCGGGGCGGAGTCTTACCTCGCTGGTGTGTGTGTGTGTGTGTGTGTGTGT 7290
 QY 8156 ctgtgaataaaatct 8198
 Db 7291 CTGTGATTAATATCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7333

RESULT 15
 AR130474 AR130474 7333 bp DNA PAT 16-MAY-2001
 LOCUS Sequence 2 from patent US 6190861.
 DEFINITION AR130474
 ACCESSION AR130474
 VERSION AR130474.1 GI:14118799
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 7333)
 AUTHORS Fishman,J.A.
 TITLE Molecular sequences of swine retroviruses method of using
 JOURNAL Patent: US 6190861-A 2 20-FEB-2001;
 FEATURES
 source location/Qualifiers
 1..7333
 BASE COUNT 1984 a 1788 c 1885 g 1676 t
 ORIGIN

Query Match 66.9%; Score 5489.8; DB 6; Length 7333;
 Best Local Similarity 83.0%; Pred. No. 0;
 Matches 6856; Conservative 0; Mismatches 402; Indels 1005; Gaps 12;

QY 1 gtgtgtgacgacgt 60
 Db 11 GT 70
 QY 61 aagacgctctgt 120
 Db 71 AAGACCGCTTGT 130
 QY 121 tgtcttcttactgt 180
 Db 131 TGTCTTCTTACTGT 190
 QY 181 caccgagagacgagctgt 240
 Db 191 CACCGAGAGACGAGCTGT 250
 QY 241 ggcgtctgt 300
 Db 251 GGCCTCTGT 310
 QY 301 ctgagacgtagagctgt 360
 Db 311 CTCAGACCGTAAAGACTGT 370
 QY 361 ccaacctgt 420
 Db 371 CCACCTGT 430
 QY 421 gtgcgtcagagagagagctgt 480

Db 431 GTCGTCAGAGAGACCGAGTCTCTGTTGTAAGCCAAAGCTTCCCTCCGCGCGCTCG 490
 QY 481 actcttttgcctgttggaaagacgagacggttgcgtgtgtctgcatctgtt 540
 Db 491 ACTCTTTTGCCTGTGTTGGAAAGACGCGGAGCGGTGCGGTGTGTGTGTGTGTGT 550
 QY 541 ctgtctgtgtgtctgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 600
 Db 551 CTGT 610
 QY 601 tgaactaccccccttgaatttgaacttcgaacttcgaacttcgaacttcgaacttc 660
 Db 611 TGACTACCCCCCTTAGTTTGCATCTGCACCTTGTGACTGAGTTAGATTCAGAGCT 670
 QY 661 atttgaacttcgaacttcgaacttcgaacttcgaacttcgaacttcgaacttcga 720
 Db 671 ATTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 730
 QY 721 caticgaltgttgaatgtgcatcgaacttcgaacttcgaacttcgaacttcgaact 780
 Db 731 CATTCGATGTTGATGTCATCAGAGGAGACCTTTAATCTGAAATTAATCTGCTGCT 790
 QY 781 aggcacatcatttctcagacttcgaacttcgaacttcgaacttcgaacttcgaact 840
 Db 791 AGGCATATCTTTTCAAGACTGACCCGCTCTCATCTGATCAGAGGCTTATATCTTA 850
 QY 841 cgttgcaagatttggcgaagatcttccgcatcgttggtttaaaccatgctaataa 900
 Db 851 CGTGGCAAGATTGTTGGCAAGATCTCTCGCATGGTTAAACCATGGCTTAATAACCA 910
 QY 901 gaaagccaggttcccgatcctgtctgttggagaaacaaacacacacacacacac 960
 Db 911 GAAAGCCAGGTCCCGAATCTGCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 970
 QY 961 tgaagcccttctctcgtatctaccccgagatcgaagagccgagacttcgcccga 1020
 Db 971 TCGAGCCCTC-TCCTGTATCTACCCGAGATGAGAGACCCGCGCACTTGGCGGAA 1029
 QY 1021 caacgttctcccccaccccttctcagacagaggtgtgtgtgtgtgtgtgtgtgt 1079
 Db 1030 CAACGTGTTCCTCCACCTCTTATCCAGACAGAGGTGTGTGTGTGTGTGTGTGT 1089
 QY 1080 ccttgagctcctgt 1139
 Db 1090 CCTGGAGCTCCGGT 1149
 QY 1140 ccggaacggaacagagagatctgcataataccgttcgacacacacacacacacac 1199
 Db 1150 CCGGAGCGGACAGACAGATCGCATATTACCGCTGCGCACTATAGGCCCTCCATG 1209
 QY 1200 gggggcccaattgcagccccctcagatctgcagcttcttctgcagatctcctat 1259
 Db 1210 GGGGGCCAAATTGCAGCCCTCCAGATATGCGCTTTCTTCTGCGAGATCTCATNA 1269
 QY 1260 aaaactaacatccccccttctcgaagatccccaacgctcgaaggttctgtgtgt 1319
 Db 1270 AAAACTAACCATCCCTCTTCTCGGAGATCCCAACGCCCTCAOGGGTGTGTGA 1329
 QY 1320 cttatgttctcaccagcctacttggatgttgcacagcttcgctgcagacacttc 1379
 Db 1330 CTTATGTTCTCTCACACCTCACTTGGATGTGTCAACAGCTGCTGCACACTCTTC 1389
 QY 1380 acaacggaagagcgaagagatctgtttagagcttgaagaaatgttccctggggcg 1439
 Db 1390 ACAACCGAGAGAGAGAGATCTGTTAGAGGCTAAAAAAATGTTCTCGGGGCGAC 1449
 QY 1440 gggcgacccagcagttgcaaatgagatlgacatggagatttccctgcagcccggt 1499
 Db 1450 GGGCGACCCAGCAGATTGCAAAATGAGATGACATGGGATTTCCCTGACTCGCCCG 1509
 QY 1500 tgggaactaacacagcgtcgaaggtcagagagacttgaatctatcgcagacttgtg 1559
 Db 1510 TGGGACTAACACAGCGGTGAGAGGAGGAGGCTTGAATAATCTATCCGCGAGGCT 1569
 QY 1560 ggggtgtctccggcgccctcgaagacgcccactaatttggctaaagtgaggtgtg 1619
 Db 1570 GCGGGTCTCCGGGGCCCTCAAGACGCGCCACTAATTTGGCTAAGTAGAGAGGTG 1629
 QY 1620 caaggacccgaagaaactccctcgttattcttgaagagctatgaagacccctcgg 1679
 Db 1630 CAGGAGCCGAGACCACTCCCTCGTATTTCTTGAAGGCTTATGAAACCTTCAGCG 1689
 QY 1680 ttaaccccttltgaatcctcctcgaagagcccaagacccctcagcttgcctcct 1739
 Db 1690 TTACCCCTTTTGTATCTCTCACTCAGAGGCCAGAAAGCCCTCAGTGCCCTTCA 1749
 QY 1740 gggcagtcgctcgtatctcgaagaaacttcgaagacccctcgaagaggttcaag 1799
 Db 1750 GGGCATCTCGCTCTGTGATTCAGAGAGAACTTCAGAGCTGAGAGGCTTACAGAG 1809
 QY 1800 gagttcgtatctatgtgagagagcagagaggtgtatcagaagagagacagagag 1859
 Db 1810 GAGTTACGTGATCTAGTGTAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1869
 QY 1860 ggaagagaaacagagaaagaaagagagagagagagagagagagagagagagag 1919
 Db 1870 GGAAGAGAAACAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1929
 QY 1920 caagaaagaaatttgaactaagatcttgcgcaagtgtgttgaagaaagagagagag 1979
 Db 1930 CAAGAGAGAAATTTGACTAAGATCTTGGCCGCTGAGTTGAAGGAGAGAGAGAG 1989
 QY 1980 agagagagagatcttgaagaaatltagtcagagccctcagacagctcagagaa 2039
 Db 1990 AGAGAGAGAGATTTAGAGAAATTTAGTCAAGCCCTACAGAGCTCAGGAACTTGG 2049
 QY 2040 agagcccaactcgaagaaagagagagagagagagagagagagagagagagagag 2099
 Db 2050 AGAGCCCACTCTGAG 2109
 QY 2100 aactgccccaaag 2159
 Db 2110 AACTGCCCAAG 2169
 QY 2160 taaggag 2219
 Db 2170 TAGGGAG 2229
 QY 2220 caacagttgagttccctcgttgcagacagagagagagagagagagagagagag 2279
 Db 2230 CAACAGTTGAGTCTCTGTTGATACCGAGAGAGAGAGAGAGAGAGAGAGAGAG 2289
 QY 2280 ggaactaacaaagaaagaaatccctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2339
 Db 2290 GGAATACTAAAGAAAGAAATCTCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2349
 QY 2340 tgaactacccgaagaaacggttgcagttgtgtgtgtgtgtgtgtgtgtgtgtgt 2399
 Db 2350 TGGACTACCCGGAAGAAACGTTGACTTGGAGTGTGGAGGAGGCTAACCACTCTG 2409
 QY 2400 atccctgagtgccagagacccctcctcagatgaagagagagagagagagagagag 2459
 Db 2410 ATCCCTGAGTCCAGTACCTCTTCTTGAAGTGAAGTCTGACCAAGATGGAGGTCAA 2469
 QY 2460 atttcttgaagaaagaaagaaagagagagagagagagagagagagagagagag 2519
 Db 2470 ATTCTTTTAAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2529
 QY 2520 accctcgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 2579
 Db 2530 ACCCTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2589
 QY 2580 atacagtcctgt 2639
 Db 2590 ATACAGTCTCTGTTGAGAGAGTTTCCCAAGGCTGTGGAGAGAGAGAGAGAGAG 2649

Db	4002	CCCTCCCGAGCTGGTTAATGCTAATCCCTTCCAGAAATACCTCCAGAAAGAGACTAAAGGGGA	4061
QY	4860	agccaccacagcgctcctacctcgggaagctggaacttcaactggaagtaaacgcggctcaatacagga	4919
Db	4062	AGCCACCCAGGCGCTCACTGGGAAGTGGACTTCACAGAGGTAAAGCGCGCTAAATACGGA	4121
QY	4920	aacaatctcatcggtttcttgtagaacacctttaaaggatgggttagagacctatccact	4979
Db	4122	AACAAATATCTAATGGTTTTTGTGACACCTTTTCAGATGGGTTMAGGCTTATCTCT	4181
QY	4960	aagaaagagacttcaaccgctggtgctgaagaaatactcagaggaattttccaaagatt	5039
Db	4182	AAAAAAGAGACTTCAACCGGTGGTGGCTTAAGAAATATCTGAGGAATTTTTCAGAGATT	4241
QY	5040	ggaataactaagygataaaggtcaagacaatggtccagcttctgctgcaggaagtcag	5099
Db	4242	GGAATACCTAAGGTAAGTAAAGGTCAGCAAAAGTCCAGCTTTGTGGCCAGGTAAAGTCAG	4301
QY	5100	ggaactggcccaagatatctgggaattctgttgaacatcgtcatctggtgcaaaagaccacaagc	5159
Db	4302	GGACTGGCCAGATATTGGGGATTGATTGAAAATGTCATTGGCATACAGACCCCAAGC	4361
QY	5160	tcaagacaagtgtagaagagacgaa taagacattaaagaaagaccttacaatctgacaca	5219
Db	4362	TCAGACAGAGTAGAGAGAGATGAATAGACCATTAAGAAAGACCCCTTACCAATTAGACACA	4421
QY	5220	gaagctcgcattaatgatttgatgctcctcctgccttgctgtcttttaagtgtagaac	5279
Db	4422	GAGACTGGCATTAATGATGTGATGGATGGCTTCCTCCCTTGTCTTTTAAGGTAGGAGAAC	4481
QY	5280	acccttggaagcttggtgcgcgaccaccttgaaattgcttaaggaggagcccccgcttg	5339
Db	4482	ACCCTTGACAGATTGGGCTGACCCCTTAATTAATGCTCTAAGGGGAGCCCCCGCTG	4541
QY	5340	gcagaaatgcctcttgacacatagtctgtatgctgcgcttcccaagccttgctctcag	5399
Db	4542	GCAGAAATTGCTTGTGACATAGTGTGATGTGTGCTTTCCACACCTTTGTCTCTAGG	4601
QY	5400	ctcaaggcgctcgagtggtgtgaagcgagcgctggaagcagctccgggaagcctactca	5459
Db	4602	CTCAAGCGCTCAGATGGGTGAGGAGCGACGAGTGGGAAGCAGCTCCGGGAGCCACTCA	4661
QY	5460	ggaaggaacttggaagttccaatgcctccaagtttgagattcaagttaagtaagc	5519
Db	4662	GGAGGAAGCTTGCAAAGTCCACATCCCTTCCAAAGTGGAGATTCAGTTATGTTAGACGC	4721
QY	5520	caccgttcagaaaacctcggagactcggtggaagggagcctatcctgtaactcttgaacaca	5579
Db	4722	CACCGTCAGAAAGAACTCCAGAGATCGGTGAAAGGAGACCTTATCTGTACTTTTACACA	4781
QY	5580	ccaacggctgtgaagtcgaaggaatcccaactcgtgatacgtacatcccaactgaagcgg	5639
Db	4782	CCAAAGGCTTGAAGATCGAAG-----	4803
QY	5640	ggcgaccactccgacctcggggtgggagagccgaagaaagactggaatcccttaagctcgc	5699
Db	4804	-----GAATCCCTTAAACCTTCGC	4822
QY	5700	ctccatcgccctgttctcttactctaaacaatactcccaagcagtagttaaagccttat	5759
Db	4823	CTCCATTCGGGTGTTCTCTTACTCTGTGCATAAATCTCTCAAGTTAATGTTAAAGCCTTGT	4882
QY	5760	agaacagctcgaaaccccatagaaacttatacccttaactcgtgctgtatcttgaacctgtatc	5819
Db	4883	GGACAGCCCGAATCCCATAAACCTTATCTCTCACTCGTTACTGTACTCGGTAC	4942
QY	5820	gggtgtcactcgtaaatagaaactcggggtgtgtcctctagaaagcaactcgtgtgctgaact	5879
Db	4943	AGGATTTAATTAATTAACAGACATCAAGGGAGGCTCCCTTGGGAGACTGGTGCGCTGAATT	5002
QY	5880	gcaattctcgctcggattgattaaccccgcttta-----aaagacaactcccaact	5933
Db	5003	ATATGCTGCTTCGATCGATAGTATATCCCTGTGTTCAATACCAAGGCGCAACCCCGCATGT	5062

QY	5934	atctccagtagttagtggtgtctctctctcccaagg---cacagagaaagagaaatctatcttg	5993
Db	5063	ACTCCGTGCTTAACGGGTTTACGTTTGTGCCAGGACCCCAATATATGAGAAATATTGTGG	5122
QY	5991	gggtctcgggaatccctctctgtagagatgagctgtctacccccaagatgagagctg	6050
Db	5123	AAATCCTCAGGATTTCTTTTCCAGCAATGAGCTGCATAACTTCAATGATGGAAATG	5183
QY	6051	gaatatggcagatctctcccaaggacgggttaaaatctctctgttcaatccc-----	6100
Db	5183	GAAATGCCAGTCTCTCAGCAAGACAGATGAGTTACTCTTTGTTTACAAATCCATCCAG	5244
QY	6103	-----99cccgggccaagtacaaga	6122
Db	5243	TTATATACATTAATTATATGSCCATGGAGATGGAAGATTTGGCAACAGGGGTACAAA	5300
QY	6122	tgaatgaatatataaagataaagcgtctctcccacacagacttagattatctaaagata	6181
Db	5303	AAATG-TACGAATATAGCAATAAAGCGTCAATTCGTTAGACCTTAGATTTCTTAAATATTA	5361
QY	6182	gtttcacgc---aaagaaacacgaaataatctcaagtgataaagtgtatgtagcttgg	6233
Db	5362	GTTTTACGTAAGAAAGAAACAAAGAAATATCTCAAAAGTGGTAAATGGTATATCTTGGG	5422
QY	6240	gaatgatttltat---tataaggcggggagacaggtctacattcaaccattgcctta	6295
Db	5422	GAAATGCTACTATATGAGAGGCTCTGGGAGAAAGAAAGATCTGTTGCACTATTTGCCCTCA	5481
QY	6296	ggaatagaacgggagacagaaacccctctgtgcaatlggaccccgataaagtaactgtctaac	6355
Db	5482	GAAATGAACACTCAGTGGAACTCTCGGTTCTATAGGACCAATATAGGGTTTGGCCGAAAG	5541
QY	6356	agggggcccccggcccttgagaccacagataactctgcggtgcccccaattaactgcgtgc	6415
Db	5542	AAGGACCTCAATCCAAAGA-----ACAGA	5565
QY	6416	ggcctgacaataacacagccgctgagcaacgctccacatgattgtatctctaccaaacgcg	6475
Db	5566	GGCCATCTCTTAACCCCTCTGATTATCAATCAACACTCTGGATCAGTCCCACTGA-----	5620
QY	6476	ctagaacatccccaagtgcttcctgtttaaagacagacagagactctcagttccatccaaag	6535
Db	5621	-----GCCATACATCATATTTAAACAGAGGGCGAAACTTTTATAGCCATCACCAGG	5670
QY	6536	gagcttccaagacataactccacgcgactgacacgtctccactctctctctgttgcttgctc	6595
Db	5671	GAGCTTTTCAAGGCTCTTAACCTCACGACACTCCAGAGGCTCACTCTTGTGGCTTGTGCT	5730
QY	6596	tatcccaaggccctccctatatgagggagtgctctaaagaagaatctcaatgtagaca	6655
Db	5731	TAGCTTGGGGCCACTTACTATGAGGGAATGGCTTAGAGAGAGGAAATTCATATGTAACA	5790
QY	6656	aagagcatagaanaatcaatgtacatlgggggtcccgaaataaagtctaacccatcaagtctt	6715
Db	5791	AGGAACATAGAGACCATGTATACATGGGGATCCCAAAATTAAGCTTACTGAGGTTT	5850
QY	6716	ccggagagggaacatgatatgagaaagctcccccataccacaacaaactttgcatagta	6775
Db	5851	CTGGAAAGGCACTGTGATAGGAGTGGGTCCCCCATCCACCAACAACTTTGTATACACA	5910
QY	6776	ctgtggttctatgacgagccctcagaaatcagtatctagtaacctgtgtlataacagtggt	6835
Db	5911	CTGAAGGCTTTAATCGAACCCTGTAGAGTCAATATCTGTGTACTGTGTTATGACAGTGTG	5970
QY	6836	gggacatgaataacttggtttaacccctgtgttccaactcagltcttcaaccaatccaaag	6895
Db	5971	GGGCAATATATACGTATTAACCCCTTGTTGTTCACACTTGGTTTTCACCAAACTAAG	6030
QY	6896	attctgtgtcatgtgttccaatctgtcccccaggttaactacacatcttgaagaagtgtgc	6955
Db	6031	ACTTTGCGTTATGTGTCAAATATGTTCCCGGGGTGTACTACTATGTCGGAAGAGAGTGC	6090

Mon Feb 25 07:44:11 2002

us-09-171-553b-3.feb22std.rge

Page 66

/translation="MHPILSRRLPTRGGEPRKRLIPISFASIAMFLTLTTTPQASSK
RLIDSNHRRLPLSTLWLTIDPTGVTNVSTRGVAPRGTMPELHRLILNPAKSTP
PNLVSIFCCPCEKREKCYGSGSEFCRWSGDSNDGMKPI SLODVKVSPN
SGPKYKMKLTKDKSCSPSDLYLKTSETERKSGKSKVKWELGNSFLYGGAG
STLTLRLIETGEPYPMGDKVLAEGEPYKPKHLPVPLPOLSTLRDPTQPPNS
TGLIPTNTPRNSPGVPYKTGORLPSLIQGAFOA INSDPATISSCMCLSSGPPYE
GMAKERKNVTKRHNOCTGWSRRKLTILEVSGGTCIGKAPSHOHLCSTVVEQA
SENOYLPGYNRMWACNTGLTPCYSTVFNSKDFCVAVOIVPRVYHHPREVLEEDY
YRNRPKREPVSLTLAVMLIGTAVGVTGALITGPQOLEKGLIHAAMEDLRA
LKEVSNLDESLSLEVLONRGLDLFLREGICLALKKECCFVYDHSGLRDSM
NKLRRKLERRERREADQGEFGEFNMSPMTTLLSALTGPIVYLLLLLVYVPCILNR
FVAFVREVSVAOIVMLRQOYGLSGEMDL"

BASE COUNT 2168 a 2064 c 2154 g 1823 t
ORIGIN

alignment_scores:

Quality: 2801.00 Length: 524
Ratio: 5.345 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-171-553b-4 x A66553 ..

Align seg 1/1 to: A66553 from: 1 to: 8209

1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrpTh 17
|||||
588 ATGGGACAGACAGTACTACCCCTTAGTTGACTCTCGACCATTTGGAC 637
17 rGluValArgSerArgAlaHisAsnLeuSerValGlnValLysGlyP 34
|||||
638 TGAAGTTAGATCCAGGGCTCATATTGTTCAGCTTAAAGAGGGAGC 687
34 rGTrGlnThrPheCysAlaSerGluTrpProThrPheAspValGlyTrp 50
|||||
688 CTGGCAGACTTCTGCTCTCTGAAATGAGCAACATCGAGTGTGAGTGG 737
51 ProSerGluGlnThrPheAsnSerGluLeuLeuAlaValLysAlaI 67
|||||
738 CCATAGAGAGGGAGCTTTAAATTCGAAATATCTCGGCTGTTAAGGCAT 787
67 eLlePheGlnThrGlyProGlySerHisProAspGlnGluProTyrIle 84
|||||
788 CATTTTACAGACTGACCCGGCTTCATCCGATCAGAGGCCCTATATCC 837
84 eLThrTrpGlnAspLeuAlaGluAspProProTyrValLysProTrp 100
|||||
838 TTACGTGGCAGATTGGCAGAGATCTCCGCCATGGTTAAACCATMG 887
101 LeuAsnLysProArgLysProGlyProArgLLeuLeuAlaLeuGlyL 117
|||||
888 CTAATATTAACCAAGAAAGCCAGGTCGCCGAACTCTGCTCTGGAGGAA 937
117 sAsnLysHisSerAlaGluLysValGluProSerSerSerTyrLeuPro 134
|||||
938 AAACAACACTCGGCCGAAAGCGAGCCCTCTCTCTGATCTATACCC 987
134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150
|||||
988 GAGATTCAGAGAGCCGCGACTTGGCCGGAACCCCAACCTGTCCCCCAGC 1037
151 ProLeuSerSerThrGlyCysGlyGlyThrSerAlaProProGlyAl 167
|||||
1038 CCCTTATCCAGCACAGGCTCTGTGAGGGGACCTTGCCTCTCTGAGAG 1087
167 aProValValGluGlyProAlaAlaGlyThrArgSerArgArgGlyAla 184
|||||
1088 TCCGGTGGTGGAGGAGCTCTCCCGGAGATCCGAGCCGAGAGGGGCCA 1137
184 hrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrGly 200
|||||
1138 CCCCAGAGGGAGCAGACAGATGCGATATTACCGCTCGCAGCACTATGCG 1187

201 ProPrometProGlyGlnLeuGlnProLeuGlnIleTyrTrpProPhe 217
|||||
1188 CTTCCCAATGCCAGGGGGCCAAATGGACCCCTCCAGATTTGGCCCTTTTC 1237
217 rSerAlaAspLeuTyrAsnTrpLysThrAsnHisProProPheSerGlu 234
|||||
1238 TTCTCAGACTCTCTATATTGGAAACCTAACCATTCCTTCTTCGAGAG 1287
234 sProGlnArgLeuThrGlyLeuValGluSerLeuMetPheSerHisGln 250
|||||
1288 ATCCCAACAGCCTCAGCGGCTTGGAGTCCCTTATGTCTCTCACACAG 1337
251 ProThrTrpAspAspCysGlnGlnLeuLeuGlnThrLeuPheThrTrp 267
|||||
1338 CCTACTTGGAGTGAATGTCACACAGCTGCTGACAGACTCTTCCACACCGA 1387
267 uGluArgGluArgIleLeuLeuGlnAlaArgLysAsnValProGlyAla 284
|||||
1388 GGAGCGAGAGAGAAATTGTTAGAGGCTAGAAAAAATGTTCTCTGGGCCG 1437
284 sPGLYArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLeu 300
|||||
1438 ACGGGCAACCCACGACGCTTGCAAAATGAGATTGACATGGGATTTCCCTTG 1487
301 ThrArgProGlyTrpAspTyrAsnThrAlaGlnGlyArgGluSerLeu 317
|||||
1488 ACTGCCCCGGTTGGACTACACACGCGCTGAAGTAGGAGGAGACTTGAA 1537
317 sLeuTyrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArg 334
|||||
1538 AATCTATGCGCAGGCTGCTGCGCGGCTCCGGGGGCGCTCAAGACGG 1587
334 rOThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGluPro 350
|||||
1588 CCACATATTGGCTTAAGTAAGAGAGTATGACGAGACCGCAACCACT 1637
351 ProSerValPheLeuGluArgLeuMetGluAlaPheArgArgPheThr 367
|||||
1638 CCGCTGATATTCTTAGAGAGCTCATGGAAGCTTAAAGCGGTTCAACCC 1687
367 oPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 384
|||||
1688 TTTTGATCTCACTCAGAGAGCCAGAAACCTCAGTGGCCCTGCGCTTCA 1737
384 LeGlyInSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGly 400
|||||
1738 TTGGGCAAGTCGGCTCTGGATATCAGGAAGAACTTCAGAGACTGGAAGG 1787
401 LeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluLysVal 417
|||||
1788 TTACAGAGAGCTGAGTGTAGTGTAGTACAGAGAGGAGAGAGAGTGTGA 1837
417 rTyrArgGlyGluThrGlnGluGluLysGluGlnArgLysGluVal 434
|||||
1838 TTACAAAGAGGAGACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1887
434 rGluGluArgGluArgArgArgAspArgArgGlnGluLysAsnLeuThr 450
|||||
1888 GAGAAAG 1937
451 LysIleLeuAlaAlaValValGluLysSerSerArgGluArgGluArg 467
|||||
1938 AAGATCTTGGCCGCGAGTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1987
467 gAspPheArgLysIleArgSerGlyProArgGlnSerGlyAsnLeuGly 484
|||||
1988 AGATTTTGAAGAAATTTAGTACAGGCCCTAGACAGTCAAGGAACCTGGGCA 2037
484 sNArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGly 500
|||||
2038 ATAGACCCCACTGCACAGAGCAGATGCGGTATTTGTAAGAGAGAGAGAG 2087


```

501 HistpAlaArgAsnCysProLysGlyAlaValGlyProLysValle 517
|||||
2088 CACTGGGCAAGAACTGCCCCAAGAAAGGAAACAAAGCCGAGGTCTCT 2137
517 uAlaLeuGluGluAspLysAsp 524
|||||
2138 AGCTTAGAAGAAATAAAGAT 2159

seq_name: gb_pat:A66552

seq_documentation_block: 8196 bp DNA PAT 29-MAR-1999
LOCUS A66552
DEFINITION Sequence 2 from Patent W09740167.
ACCESSION A66552
VERSION A66552.1 GI:4538105
KEYWORDS
SOURCE
ORGANISM
porcine endogenous retrovirus.
Virus; Retroid viruses; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
REFERENCE
1 (bases 1 to 8196)
AUTHORS Galbraith,D.N., Haworth,C., Lees,G.M. and Smith,K.T.
TITLE PORCINE RETROVIRUS
JOURNAL Patent: WO 9740167-A 2 30-OCT-1997;
Q ONE BIOTECH LTD (GB)
FEATURES
source
1. 8196
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
BASE COUNT 2165 a 2061 c 2146 g 1820 t 4 others
ORIGIN

alignment_scores:
Quality: 2774.00 Length: 524
Ratio: 5.304 Gaps: 0
Percent Similarity: 99.809 Percent Identity: 99.809

alignment_block:
US-09-171-553b-4 x A66552
Align seg 1/1 to: A66552 from: 1 to: 8196

1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisTrpH 17
|||||
576 ATGGGACAGACGACTACCCCTTAGTTGACTCTGCACCATTTGGAC 625
17 rGluValArgSerArgAlaHisAsnLeuSerValGlnValLysGlyP 34
|||||
626 TGAAGTTAGAACAGGCGCTCATATTGTCAGTTCAAGTTAAGAGGAC 675
34 rotPpGlnThrPheCysAlaSerCylurProrPheAspValGlyTrp 50
|||||
676 CTTGGCAGACTTCTGCTGCTCAATGCGCAACATTGCGATGGTGGATGG 725
51 ProSerGluGlyThrPheAsnSerGluIleLeuAlaValLysAlaIle 67
|||||
726 CCATCAGAGGAGGACCTTAATTCGAATATATCTGCTGTAAAGCAAT 775
67 eAlaPheGlnThrGlyProGlySerHisProAspGlnGluProTyrIle 84
|||||
776 CATTTTCAGACTGGACCGCGCTCTCATCCGATCAGAGAGCCCTATATCC 825
84 eutThrPpGlnAspLeuAlaGluAspPropPheProTyrValLysProTyr 100
|||||
826 TTACGTGGCAGATTGTGGCAGAGATCTCTCCGCTATGAGGTTAAACCATGG 875
101 LeuAsnLysProArgLysProGlyProArgLleLeuAlaLeuGlyGly 117
|||||
876 CTAATTAACCAAGAAAGCCAGGTCCTCCGATCTGCGCTTGGAGAGAA 925
117 sAsnLysHisSerAlaGluLysValGluProSerSerSerTyrIleuPro 134
|||||

926 AAACAAACACTCGGCGCAAAAAAGTCGAGCCCTTCTCTGATCTACCCC 975
134 rGaAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150
|||||
976 GAGATCGAGGAGCCCGGACTTGGCCGGAACCCCAACCTGTCTCCCAACC 1025
151 ProLeuSerSerThrGlyCysCysGlyGlyIleThrSerAlaProProGlyAla 167
|||||
1026 CCCTTATCCACACAGGAGTGTGTGAGGAGACCTCTGCGCTCTGGAGGC 1075
167 aProValValGluGlyProAlaAlaGlyThrArgSerArgArgLysAla 184
|||||
1076 TCCGCTGTGGAGGAGGACCTCTGCGGAGACCTCGAGGAGGAGGAGGAGCA 1125
184 hrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrGly 200
|||||
1126 CCCCAGGAGGAGCAGAGAGATCCGGAATATACCGCTCGGACCTATGCG 1175
201 ProprometProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPheSe 217
|||||
1176 CCTCCATGCCAGGAGGCGCAATTCGAGCCCTCCAGTATGGCCTTTTC 1225
217 rSerAlaAspLeuTyrAsnTrpLysThrAsnHisProProPheSerGlu 234
|||||
1226 TTCTGCAGATCTCTAATAATGGAAACTAACCATCCCTTCTCGGAGG 1275
234 sProGluArgLeuThrGlyLeuValGluSerLeuMetPheSerHisGln 250
|||||
1276 ATCCCAACAGCCTCACGGGGTTGGAGTCCCTTATTTCTCTCACAGC 1325
251 ProThrTrpAspAspCysGlnGlnLeuGlnThrLeuPheThrThrGly 267
|||||
1326 CCTACTGGAGATGATTGCAACAGCTGCTGCAGACACTTCCACAAACGA 1375
267 uGluArgGluArgLleLeuLeuGluAlaArgLysAsnValProGlyAla 284
|||||
1376 GGAGGAGAGAGAAATTCGTTAGAGCTTAGAATAATGTTCTCGGGGCGG 1425
284 sPglYArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLeu 300
|||||
1426 ACGGGGAGCCACGCGAGTTCGCAAAATGAGATGAGATGGATTTCCCTTG 1475
301 ThrArgProGlyTyrPaspTyrAsnThrAlaGluGlyArgGlySerLeu 317
|||||
1476 ACTCGCCCGGTTGGAGACTACAAACAGCGCTGAAGTAGGAGAGAGCTTGA 1525
317 sIleTyrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArgP 334
|||||
1526 AATCTATCCCGAGGCTGTGTGGCGGTCTCCGGGGCGGCTCAAGACGGC 1575
334 rThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGluPro 350
|||||
1576 CCACATAATTGGCTTAGTAGAGAGAGATGATCAGAGGACCGAAGCAACT 1625
351 ProSerValPheLeuGluArgLeuMetGluAlaPheArgArgPheThrPr 367
|||||
1626 CCCTCGGATTTCTTGAAGGCTCATGAGAGAGCCCTTACGGCGGTTACCCC 1675
367 oPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPheI 384
|||||
1676 TTTTGATCTCACTCTAGAGGCCAGCAAAAGCCTACAGTGGCGGCTTCA 1725
1676 eAlaGlnSerAlaLeuAspLleArgLysLysLeuGlnArgLeuGly 400
|||||
1726 TTGGCAGAGTGGCTGTGATATCAGAGAAACCTTACGAGACTGGAGG 1775
401 LeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluLysValTy 417
|||||
1776 TTACAGGAGGCTAGTACGTGATCTAGTAGAGAGAGAGAGAGAGGTGA 1825
417 rTyrArgArgGluThrGluGluLysGluGlnArgLysGluLysGlu 434
|||||
1826 TTACAGAGAGGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1875

```

```

434  rgcIugluaragcIugluarargararsparargInoIuIuysnLeuThr 450
      |||||
1876  GAGAAAGAAAGGAGGAGAAAGCGATGATAGACGGCAAGAGAAATTTGACT 1925
451  LysIleLeuAlaIaIaValValgIuIyLysSerSerArgIaIaArgIuIur 467
      |||||
1926  AAGATCTTGCGCCCACTGGTTCAGAGGGAAGACAGCAGGGAAGAGACAG 1975
467  gAspPheargLysIleargSerGlyProargInSerGlyAsnLeuGlyA 484
      |||||
1976  AGATTTTAAAGAAATTAAGTCAAGCCCTTAGACAGTCAAGGAACCTGGGCA 2025
484  snArghrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGly 500
      |||||
2026  ATAGGACCCCACTCGCAAGGACCAAGTGTGCTATTGTMAAGMAAAAGA 2075
501  HisTrpAlaArgAsnCysProLysLysGlyAsnLysGlyProLysValle 517
      |||||
2076  CACTGGGCGAAGAACTGCCCAAGAGGAGGAAACAAAGACCAAGACTC. CT 2124
517  uAlaLeuGluGluAspLysAsp 524
      |||||
2125  ACCTCTAGAGAGAAAGATTAAGAT 2146

```

seq. documentation block:	3025 bp	DNA	MM	27-APR-2000
LOCUS	AF147808			
DEFINITION	Sus scrofa domestica porcine endogenous retrovirus type C gag-pol precursor, gene, partial cds.			
ACCESSION	AF147808			
VERSION	AF147808.1	GI:7650139		
KEYWORDS	.			
SOURCE	domestic pig.			
ORGANISM	Sus scrofa domestica			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
AUTHORS	1 (bases 1 to 3025)			
TITLE	Blusch,J.H., Seelmeier,S. and von der Helm,K. Evidence for extreme sequence homogeneity of the PERV A/B 5' LTR/leader/gag/prot region by predicted targeted amplification from a Du x DL pig			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 3025)			
AUTHORS	Blusch,J.H.			
TITLE	Direct Submission			
JOURNAL	Submitted (03-MAY-1999) Virology, Max-von-Pettenkofer-Institute, Pettenkofer-Str. 9a, Munich 80336, Germany			
FEATURES	Location/Qualifiers			
source	1..3025			
	/organism="Sus scrofa domestica"			
	/strain="bred Duroc x German Landbreed, Du x DL3791"			
	/sub_species="domestica"			
	/db_xref="taxon:9825"			
	1..3025			
	/note="PERV A/B; porcine endogenous retrovirus type C"			
	/rpt_family="perv"			
	/rpt_type="dispersed			
	587..605			
	/note="primer binding site for tRNA-Gly"			
	1032..>3025			
	/note="gag protein produced by proteolytic cleavage of the mature protease out of the precursor"			
	/codon_start=1			
	/transl_except="(pos:2604..2606,aa:Gln)			
	/product="gag-pol precursor"			
	/protein_id="AA65926.1"			
	/db_xref="GI:7650140"			
	/translation="MGQIVTPLSLDHWTEVRSRAHNLISVQKKGPMOTFCASEMP TFPVGSECTENSETIILAVKAIIFQTGSPHDDPEYILTWOLADPPWPKMILN KPKRPGRIILTEKKNHSAKEVPSRYPIETPEPTWPPQPPPPYPAQCAVRC PSAPGPAVVEGPAAGTBSRGATPERIDETAILPRTYGPMMGGLOLDTWPFSS			
primer_bind				
CDS				

mat_peptide		mat_peptide		mat_peptide	
2604...>3025		2604...>3025		2604...>3025	
/product="aspartic protease"		/product="aspartic protease"		/product="aspartic protease"	
BASE COUNT		BASE COUNT		BASE COUNT	
ORIGIN		ORIGIN		ORIGIN	
743	a	749	c	849	g
684	t				
<p>alignment_scores:</p> <p>Quality: 2761.00 Length: 525</p> <p>Ratio: 5.269 Gaps: 1</p> <p>Percent Similarity: 99.810 Percent Identity: 99.810</p>					
<p>alignment_block:</p> <p>US-09-171-553B-4 x AF147808 ...</p>					
<p>Align seg 1/1 to: AF147808 from: 1 to: 3025</p>					
1	MetGlyInThrValThrtThrProLeuSerLeuThrLeuAspHisTrpTh	17	rgluValArgSerArgAlaHisAsnLeuSerValGlnValLysLysGlyP	34	roTrpGlnThrPheCysAlaSerGluTrpProThrPheAspValGlyTrp
1032	ATGGGACACAGACAGACGACCCCTTACGTTTGACTCTCGACCATTTGGAC	1082	TGAAGTTAGATCCAGGGCTCTATATTGTTCAGTTCAGGTTAAAGAGGAC	1132	CTTGCGACACTTCTGTGCTCGCTCGATGGCCAACTTCGATGTTGATGAGG
1182	CCATCAAGAGGGACCTTTAATTCGAAATTAATCTCGGCTGTTAAGGCAAT	1282	TTACGTCGGCAAGATTGGCGAAGAAATCTCCGCGCATGGCTTAAACCATGG	1332	CTTAATTAACCAAGAAAGCCAGGCTCCCGAAATCTCGGCTTGGAGAGAA
1382	CAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	1432	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	1482	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC
1532	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	1582	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	1632	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC
1682	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	1732	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	1782	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC
1832	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	1882	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	1932	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC
1982	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	2032	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	2082	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC
2132	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	2182	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	2232	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC
2282	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	2332	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	2382	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC
2432	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	2482	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	2532	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC
2582	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	2632	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	2682	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC
2732	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	2782	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	2832	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC
2882	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	2932	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	2982	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC
3032	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	3082	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	3132	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC
3182	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	3232	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	3282	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC
3332	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	3382	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	3432	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC
3482	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	3532	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	3582	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC
3632	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	3682	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	3732	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC
3782	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	3832	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	3882	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC
3932	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	3982	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	4032	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC
4082	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	4132	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	4182	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC
4232	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	4282	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	4332	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC
4382	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	4432	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	4482	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC
4532	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	4582	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	4632	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC
4682	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	4732	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	4782	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC
4832	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	4882	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC	4932	GAATTTTTCAGACTGGACCCGCGCTCTCATCTGATAGAGACCCATATACC
4982	GAATTTTTCAGACTGG				

200 yProprometProglyglyInLeuGlnProLeuGlnTyrTrpProPhe 217
 1631 CCTCCCATGCGGGGGCCAAATTCAGCCCTCAGTATGCGCCCTTTT 1680
 217 eSerAlaAspLeuYrAsnTrpIysThrAsnHisProProPheSerGlu 233
 1681 CTTCTGACAGATCTATATATGAAACCTAACCAATCCCTTCTCGAG 1730
 234 AspProGlnArgLeuThrGlyLeuValGlnSerLeuMetPheSerHisG 250
 1731 GATCCCAACGCGCTCAGGGGTTGGTGAAGTCCCTATATGCTCTCACA 1780
 250 nProThrTrpAspAspCysGlnGlnLeuLeuGlnThrLeuPheThrThg 267
 1781 GCCTACTTGGAGATGATGTCACAGCTGCTGACAGACACTTTCACACG 1830
 267 InGluArgGluArgGlyLeuLeuGlnValArgGlyAsnValProGlyAla 283
 1831 AGGAGGAGAGAGAAATTTGTTAGAGGCTAGAAAAATGTTCTCGGGCC 1880
 284 AspGlyArgProThrGlnLeuGlnAsnGlnIleAspMetGlyPheProle 300
 1881 GACGGGGGACCCAGCAGTGCATAATGAGATGACATGGGATTTCCCTT 1930
 300 uThrArgProGlyTyrAspTyrAsnThrArgGluArgGlnSerLeu 317
 1931 GACTCGCCCGGTTGGACTACACACGGCTGAGGTAGGAGAGCTTGA 1980
 317 ySileYrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArg 333
 1981 AATCTCTATGCCAGGCTGTGGGGGTCTCCGGGGCGCTCACAAGACGG 2030
 334 ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGluPr 350
 2031 CCCACTAATTGGCTAAGTAGAGAGGTATGATCAGAGGACCGACGAACC 2080
 350 OProSerValPheLeuGlnArgLeuMetGlnAlaPheArgArgPheThrP 367
 2081 TCCTCGGATTTCTTTCAGAGGCTCAGAGGACCTTCAGCGGTTCAACC 2130
 367 ropPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 383
 2131 CTTTGTATCTACTCTCGAGGCCAGAAAGCTCAGTGGCCCTGGCCCTTC 2180
 384 IleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGln 400
 2181 ATTGGCAGTCGCTCGATATCAGAAAGAACTTCAGAGACTGGAGAGG 2230
 400 yLeuGlnGluAlaGlnLeuArgAspLeuValArgGluAlaGlnLysValT 417
 2231 GTTACAGAGAGCTGAGTTACCTGATCTAGTCAGAGAGCAGAGAGAGTGT 2280
 417 yTrpYrArgArgGluThrGluGlnLysGlnArgGlyGlnLysGln 433
 2281 ATTACAGAGAGGAGACAGAAAGAGAGAGAACAGAAAAAGAAAGAGAG 2330
 434 ArgGlnGluArgGlnGluArgArgAspArgArgGlnGlnLysAsnLeuTh 450
 2331 AGAGAGAGAGAGAGAGAGAGAGATGAGAGGAGAGAGAGAGAGATTGGAC 2380
 450 fLysIleLeuAlaAlaValAlaGlnLysSerSerArgGlnArgGln 467
 2381 TAAAGTCTTGGCCGAGTGTGAAGGAGAGAGAGAGAGAGAGAGAGAGA 2430
 467 rGaspPheArgLysIleArgSerGlyProArgGlnSerGlyAsnLeuGly 483
 2431 GAGATTTTAGGAAAAATTAGTCAGGCCCTAGACAGTCAAGGAGACCTGGGC 2480
 484 AsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGlnLysG 500
 2481 AATAGGACCCCACTGACAGAGAGACAGTGTGCTATTTTAAAGAAAAAGG 2530

500 yHisTrpAlaArgAsnCysProLysGlyAlaGlnLysGlyProLysVal 517
 2531 AACTGGGCGACAGAGAACTGCCCAAGAGAGAGAAACAAAGAGAGAGTCC 2580
 517 euAlaLeuGlnGluAspLysAsp 524
 2581 TAGCTCTGAGAGAGATAAAGAT 2603
 seq_name: gb_v1:PEIN33817
 seq_documentation_block: 8849 bp DNA VRL 07-JAN-2000
 LOCUS PEIN33817
 DEFINITION Porcine endogenous retrovirus type C proviral gag, pol and env
 genes and LTR (class A, clone 42).
 ACCESSION AJ133817
 VERSION AJ133817.1 GI:6688947
 KEYWORDS env gene; env protein; gag gene; gag protein; pol gene; pol
 protein.
 SOURCE porcine endogenous retrovirus.
 ORGANISM porcine endogenous retrovirus
 Viruses; Retroviral viruses; Retroviridae; Mammalian type C
 retroviruses; 1-Mammalian type C virus group.
 1 (bases 1 to 8849)
 Czauderna,E., Fischer,N., Bolter,K., Krach,U., Kurth,R. and
 Toenjes,R.R.
 Molecular Characterization of Human-tropic and
 Replication-competent Porcine Endogenous Retroviruses
 Unpublished
 2 (bases 1 to 8849)
 Toenjes,R.R.
 Direct Submission
 Submitted (04-MAY-1999) Toenjes R.R., Medical Biotechnology,
 Paul-Ehrlich-Institut, Paul-Ehrlich-Strasse 51-59, Hessen, D-63225
 Langen, GERMANY
 FEATURES
 source
 1..8849
 location/Qualifiers
 /organism="porcine endogenous retrovirus"
 /strain="type C"
 /db_xref="taxon:61673"
 /clone="42"
 /note="grown in PK15 (ATCC No. CCL-33) cell line"
 1..668
 506
 /note="cap site"
 671..688
 /note="q1y4"
 1115..2686
 /gene="gag"
 1115..2686
 /gene="gag"
 /function="retroviral capsid"
 /product="gag protein"
 join(1115..2686,2690..6271,6150..8132)
 /note="viral polypeptide"
 /codon_start=1
 /protein_id="CAB65340.1"
 /db_xref="GI:6688948"
 /translation="MCGTIVTPISLTLDHMTVRSRAHNLISVOYKKGPMOTPCASEMP
 TFDVGSEGEFNSSEILAVKAILFTQPGSHPOEPIITLWDLADPPPPVAKSEPN
 PKRPGRPTILGKRNKHSKAEVSPSPRIYDEIETPTEBPQVPPVPAQAVG
 PSAPPGATVAVGCPAAGTGRSRGATPERTDELAFLPLRTYGPMPGGLOLPMPFS
 ADLNNKTNHPSESDPQRLTGLVESLMSISQPTWDDCOLOLITLTTEPERITLBA
 RKNVAGAGRTQDNEIDMGEPITRFGMDYNTKAGRESLKITRQALVAGLRGRSRP
 TNLAKVRVMQGNPEPSVFLERLMEAFRRTPDPYSEAKASVALATGOSALDIR
 KKLRLGLOLAELRDILVREAEKYVYRREDEEKEREERERERRORRDEKNTL
 KILAVVGKSSRREREDFKIRSGPQSGMLGNRPLDKDOCAVCEKGMANRCPK
 KGNKPTALAEIDKGRGSDPLPEPPVTAKVGQVEFLYDGAHSLYLDPLGLL
 KEKSKWVGASQKQYQWTPRTYDIDGAGRVTHSELVIECPVPLGRDLITKGAQI
 SFEQGREPVSNKRPITYLTIDLDERTLRSPOYKPPRODIOISMLEOPQAAETRAGM
 LAKQVPOVDILKASAPSVSRQYPLSRHAREGLMPHYQRLIQGILYVQSPNPNPL
 LPVRPGTNDYRPVODLREVNKRVDIHPVNPYNLLSLAPERRNNYTVLDLDAF
 CLRLHPTSQPLFAFEWRDPTGRTGQLTWRLPQGFKNSTPIFDEALHRLDANRKH

```
POVTLQVVDLLAGATKQDCLEGTAKALLLESLDGLRASAKKAQICRREVTYLGYS
LGGGKRWLHAKRTKTVVOIPAPTTAKOVREFLGTAGCRIMIPGFALAPLYLITE
KGEESMAHEKADPAIKKALLAPALAPDPTAFLLYDERGVKAVGVTOTLIGM
RRPVAIYSKLDIYVAGSGWPCIKAAIAAAILVKKADLITGONTTVIAPHALENTIVO
PPDRMNTNARMTYHOSLLTERVTEAPPALNPATLIPETDEPVTCHQDLIEEIS
VRKDLTDILPGLGEYLTWPTGSSYVVGKRRAGAAYVDGTTTIVASSLPSTISQKAE
LMALTOALRLAEGSINITYDSRYAFATAHGAIVYKQGLITSAGRIKKEBISL
LEALHLEKRLAI THCPGHQKAKDLISGNOMADAVAAQAQAVNLPI IETPKAPBR
ROYTLEDMOEIKRIDOSETPEGTCTYSGKEILPHKEGLEYVOQILRLHGLHIHQ
QVLRTPSYHVLRLPGVADSVYKHCPCOLVANSPRIPGKRLGSHGAWHEVDPE
VKPAKYGNKYLVLFVDPEFSGVEAYPTKKESTYVAKLIEETEPREGICVSDNG
PAPVAOVSQGLALIGIDMKLHCAROSSGOVBRMRTKETLTKITTTGCTINDMA
LPLVLFVRVNTPGORGLTPEYELLYGSPPLVELIASVSHADVLISQLEFRLALEWY
RQAKMOLRENTYSGEDLOVPHRFQVDSVYVRHRAGNLETWKGYLVLITPTAVY
KVERIPTWHAISHVYPAPPDGSKWKAETENPKLRLHRVYVYVNNSSMHPETLSR
HLPIRGKRPKRLKIPLSFASIAMFLTSLTPOVYKRLVDSPNHKLSTLTLTDSG
TGININSTOGAPLGTWPELYVCLRSVIGLNDQAPPLVRLAYGYVCGPDPNDE
YCNODEPFCOMSCISNDGNMKMYPGSDRYSYVNNPTSYNPNYHGRKMDQ
OROKDVRNKIOTISCHSIDLYLKIPTFEKQKORNTQWNGISGIVYIGSGRKKGS
VLTIRLLETOMEPVPAIGFNKGLAEGPPIQORPSPNSDYNTTSGSYPTENITTI
KTGAKLPSLIQAGAOALNSTTPEATSSCWLCLASGPYRBMARGFNTKEDRQC
TWSQNKLTLEVSGKTCIGWVOPSHOHCNHTAEAFNRTSESOYLVPGYDRWACNT
GLTPECVSTLVENOTKDCVWVOIYPRVYYPEKAVLDKTYRYRNRPKREPISLTIAVM
LGLGVAGVGEGTAAITGPOLBKGLSNLHRTVTEDLQLEKSYNLESLSLSEY
VLONRRGLDLFLKEGGLCYALKRECFYVDHSGAIRDSKILREKIRRREREADQ
GMPEGWNRSPPMVTTLISALPLVLLLLTLVPCILNRFVAFRERSAVQIMVLR
OQTOGLLSQGETLD"
mat_peptide
/gene="pol"
/feature="retroviral protease and polymerase"
/product="Pol protein"
2690..6271
/gene="pol"
6150..8129
mat_peptide
/gene="env"
/function="retroviral envelope"
/product="Env protein"
6150..8129
gene
/gene="env"
8182..8849
LTR
8742..8747
polyA_signal
2355 a 2191 c 2281 g 2022 t
BASE COUNT
ORIGIN
alignment_scores:
Quality: 2761.00 Length: 525
Ratio: 5.269 Gaps: 1
Percent Similarity: 99.810 Percent Identity: 99.810
alignment_block:
us-09-171-553b-4 x PEN133817 ..
Align seg 1/1 to: PEN133817 from: 1 to: 8849
1 MetGlyInThrValThrThrProLeuSerLeuThrLeuAspHisTrpTh 17
|||||
1115 ATGGAGACGACAGTACTACCCCTTAGTTGACTCTCGACCACTTGAC 1164
17 rGluValArgSerArgAlaHisAsnLeuSerValGlnValLysLysGly 34
|||||
1165 TGAAGTAGATCCAGGCGCTCATATTGTCAGTTCAGGTTAAGAAGGAC 1214
34 rOTPrGInThrPhcCysAlaSerLutrrProThrPhcAspValGlyTrp 50
|||||
1215 CTTGGCAGACTTCTGTGCTCGAATGAGCCACATTCGATGTTGGATGG 1264
51 ProSerGluGlyThrPhcAsnSerGluIleIleLeuAlaValLysAlaI 67
|||||
1265 CCATCAGACAGGAGACCTTAATTCGAAATATCTGCTTTAAGGCAAT 1314
67 eIlePhcGInThrGlyProGlySerHisProAspGlnGluProTyrIle 84
|||||
1315 CATTTTCAGACTGAGACCCGGCTCTCATCTGATCAGGAGGCCATATATCC 1364
84 eutThrTrgGlnAspLeuAlaGluAspProProProTrpValLysProTrp 100
|||||
1365 TTACGTGGCAGATTGGCAGAGATCCCTCCGCACTGGGTAAACCATGG 1414
101 LeuAsnLysProArgLysProGlyProArgGluLeuAlaLeuGlyGly 117
|||||
1415 CTAAATTAACCAAGAAACCAAGCTCCCGAATCTCGCTTGGAGAGAA 1464
117 sAsnLysHisSerAlaGluLysValGluProSerSerSerTyrLeuPro 134
|||||
1465 AAACAAACACTCGGCCGAAAGTCAAGCCCTC.TCCTCGATATACCCC 1513
134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150
|||||
1514 GAGATCGAGGAGCCGCGACTTGGCCGGAACCCCAACCTGTCCCCAC 1563
151 ProLeuSerSerThrGlyCysCysGlyGlyThr.SerAlaProProGly 167
|||||
1564 CCCTTATCCACACAGGGTGTGTGAGGGAGACCTCTGCCCCCTCTGGAG 1613
167 lArProValValGluGlyProAlaAlaGlyThrArgSerArgArgAla 183
|||||
1614 CTCGGGTGTGGAGGAGACTGCTGCGCGACTCGGACCGGAGAGGCCGC 1663
184 ThrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrArg 200
|||||
1664 ACCCGGAGCGGACGACGAGATCGGATATTCCTGCGCACCTATATGG 1713
200 yProProMetProGlyLysLysLeuGlnProLeuGlnTyrTrpProPhe 217
|||||
1714 CCTCCCATGCGACAGGGGCCAATTCAGCCCTCCAGTATATGGCCCTTT 1763
217 eSerAlaAspLeuThrAsnTrpLysThrAsnHisProProPheSerGlu 233
|||||
1764 CTCTCGAGATCTCTATTAATTTGAAAACTAACCATCCCTTCTCGGAG 1813
234 AspProGlnArgLeuThrGlyLeuValGluSerLeuMetPheSerHis 250
|||||
1814 GATCCCAACGCTCAAGGGGTGTGGAGATCCCTTATGTTCTCTACCA 1863
250 nProThrTrpAspAspCysGlnLeuLeuGlnThrLeuPheThrThrG 267
|||||
1864 GCCTACTTGGATGATGATGTCACACGCTGTCGACACTCTTCACAAAG 1913
267 lUGluArgGluArgIleLeuLeuGluAlaArgLysAsnValProGlyAla 283
|||||
1914 AGGAGGAGAGAGAAATTTGTTAGAGGCTGAGAAAAAATGTTCTCGGG 1963
284 AspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPhePro 300
|||||
1964 GACGGGCGAGCCGACGAGTTGCCAAATGAGATTGACATGGGATTCCTT 2013
300 uThrArgProGlyTyrAspPtyrAsnThrAlaGluGlyArgGluSerLeu 317
|||||
2014 GACTCCGCCCGGTTGGAGCTACACACGCGCTGAGGTAGGAGACTTGA 2063
317 ySileTyrArgGlnAlaLeuValAlaGlyLeuArgGlyLysSerArg 333
|||||
2064 AATCTATGCGCCAGGCTCTGTGTGGGGGTCTCCGGGGCCCTCAAGAG 2113
334 ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGly 350
|||||
2114 CCACATAATTTGGCTAAGTAAGAGAGCGATGACGAGGAGCCACAAAG 2163
350 oProSerValIleLeuGluArgLeuMetGluAlaPheArgArgPheThr 367
|||||
2164 TCCCTCGGATTTCTTGAAGAGCTCATGGAAGCTTCAAGGGGTTACACC 2213
367 rOPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 383
|||||
2214 CTTTGTATCTTACTTCAGAGGCCCAAGAAAGCTCTAGTGGCCCTGCTTC 2263
```

```

384 11leglyInserAlaLeuAspIleArgLysLysLeuGlnArgLeuGlnI 400
|||||
264 ATTGGCGCTGCTGTGATATCGAAGAACTTCAGAGACTGGAGAG 2313
|||||
400 yLeuGlnGlnIaGlnLeuArgAspLeuValArgGlnIaGlnLysValT 417
|||||
2314 GTTACAGAGAGCTGAGTTAGCTGATCTAGTGAGAGAGCAAGAAGTGT 2363
|||||
417 yTTrArgArgGluThrGlnGlnLysGlnGlnArgLysGlnLysGln 433
|||||
2364 ATTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2413
|||||
434 ArgGlnGlnArgGlnGlnArgArgArgArgGlnGlnLysAsnLeuTh 450
|||||
2414 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2463
|||||
450 rLysIleLeuAlaIaValaGlnGlnLysSerSerArgGlnArgGln 467
|||||
2464 TTAAGATCTTGGCGCGAGTGTGAAGGAGAGAGAGAGAGAGAGAGAG 2513
|||||
467 rGAspPheArgLysIleArgSerGlyProArgGlnSerGlyAsnLeuGly 483
|||||
2514 GAGATTTTAGGAAATAGGTGAGGCGCTAGACAGTCAAGGAGACCTGGGC 2563
|||||
484 AsnArgThrProLeuAspLysAspGlnCysAlaTrpCysLysGlnLysGln 500
|||||
2564 AATAGAGACCCCTCGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2613
|||||
500 yHISrAlaArgAsnCysProLysLysGlnAsnLysGlnProLysValL 517
|||||
2614 ACACGGGCGAAGAGAACTGCCCAAGAGAGAGAGAGAGAGAGAGAGAG 2663
|||||
517 euAlaLeuGlnGlnLysAspLysAsp 524
|||||
2664 TAGCTCTAGAGAGAGAGATTAAGAT 2686
|||||
seq_name: gb_cm:AF038601
seq_documentation_block:
LOCUS AF038601 7333 bp mRNA MM 21-MAY-1998
DEFINITION Sus scrofa porcine endogenous retrovirus ERV-PK15 mRNA, complete
sequence.
ACCESSION AF038601
VERSION AF038601.1 GI:3133304
KEYWORDS
SOURCE
ORGANISM
Sus scrofa
pig.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
REFERENCE
1 (bases 1 to 7333)
Akiyoshi,D.E., Denaro,M., Zhu,H., Greenstein,J.L., Banerjee,P. and
Fishman,J.A.
Identification of a full-length cDNA for an endogenous retrovirus
of miniature swine
J Virol. 72 (5), 4503-4507 (1998)
JOURNAL
MEDLINE
98216827
REFERENCE
2 (bases 1 to 7333)
Fishman,J.A.
Direct Submission
Submitted (16-DEC-1997) Infectious Disease Unit, Massachusetts
General Hospital, 55 Fruit Street, Boston, MA 02114, USA
JOURNAL
REFERENCE
3 (bases 1 to 7333)
Fishman,J.A.
Direct Submission
Submitted (15-MAY-1998) Infectious Disease Unit, Massachusetts
General Hospital, 55 Fruit Street, Boston, MA 02114, USA
REMARK
COMMENT
On May 16, 1998 this sequence version replaced gi:3116445.
FEATURES
Location/Qualifiers
1..7333
/organism="Sus scrofa"
/strain="miniature swine"
Source

```

```

/ab_xref="taxon:9823"
/cell_line="PK-15"
/feature="porcine endogenous retrovirus ERV-PK15; infective
virus but represents a naturally occurring transcript"
1..578
598..2172
/codon_start=1
/product="gag protein"
/protein_id="AAC16766.1"
/db_xref="GI:3116446"
/translation="MGQIVTTPSLITLDHMTVEVSRANHSVYKKGPMQTECASEMP
TFDVGMPSEGEFNSLILAVAILFEQTPGSHPDQEPYILTWDLADPPWPVAPWLN
KPRKGPRIILALGKNSHAKVEPSRIYPEIEPPWPBPQVPPPPYPAQAVRG
PSAPGAVVILGSPAGTRSRGATPERTDEIALPLRTYGPMPGGLOLQLOYPFS
ADLYMKTNNHPESPDPORLTGLVESLMEFHSHPDQDOLLOTLEETTERILLLEA
KNVAGAGRPOLONEIDMGPPILTPGQMDVNTAGRESLKTIRQALVAGIRGASRP
TNLAKVREVMGPNPEPSVLERLMEARRETPPDPSSEAKASVALAIFGOSLDIR
KILQLEGLQEAELDLVREAEKVYRRETEEEKRORKEKEEREERERDRORKNIL
KILAIVVGGKSSRERERDRFRIRSGPPOSGNLGNTPIDKDCAYCKEKGMWANNCPK
KGNKPKYLALEEDKD"
2320..3552
/feature="truncated protein"
/codon_start=1
/product="pol protein"
/protein_id="AAC16767.1"
/db_xref="GI:3133305"
/translation="MGATGCGROYPMWTRRRTVDLGVYTHSEFLVPECPVLLGRDL
TKMGNOISEEGRPSVYNNKPITVLTLDLDEYRLVSROYKPPDDISMTLEQPOAV
AETAGMGLAKQVPPPOVTLQKASATVYSROYPLSEAREGIMPHVORLIOGILVPO
SWNTNPPLLEVRKPGTNDYRPVODLEVRKROVDHPTVFNPTNLISALPEPNNVTYL
DLKDAFPLRLHPISQPLTEFEMRDPGGRGQLTWLRLPGFNKSPITIDEALHRL
ANFRIOHPQVTLQYVDLLLAGATKQDCLEGTALLLLELDLGRASAKRAQICRE
VTLYGSLRGGORMLTEARKTVVQIPAPTAQVREPLTAGRCRLMIPGFATLAAP
LYPLTKRKGGLPQGGK"
4803..6725
/feature="env protein; N terminus deleted"
/codon_start=1
/pseudo
6726..7333
BASE COUNT 1984 a 1788 c 1885 g 1676 t
ORIGIN
alignment_scores:
Quality: 2758.00 Length: 525
Ratio: 5.263 Gaps: 1
Percent Similarity: 99.810 Percent Identity: 99.619
alignment_block:
US-09-171-553B-4 x AF038601 ..
Align seg 1/1 to: AF038601 from: 1 to: 7333
1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrpTh 17
|||||
598 ATGGAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 647
|||||
17 rGluValArgSerArgLahHisAsnLeuSerValGlnValLysGlyP 34
|||||
648 TGAAGTTAGACCAAGCGCTCATATTGTGTACGTTACAGATTAGAGAGAC 697
|||||
34 rTrpGlnThrPheCysAlaSerGluTrpProThrPheAspValGlyTrp 50
|||||
698 CTGGCAGACTTCTGTGCTCTCTAATAGGCCAATTCGATGTGATGG 747
|||||
51 ProSerGlnGlyThrPheAsnSerGlnIleLeuAlaValLysAlaI 67
|||||
748 CCATCAGAGAGAGACCTTAATTCGAATAATTCCTGCTGTAGGCAAT 797
|||||
67 eLlPheGlnThrGlyProGlySerHisProAspGlnGlnProTrpIle 84
|||||
798 CATTTTCAGACTGAGACCGGCTCTCATCTGATCAGAGAGAGCCCTATATCC 847

```

```

1747 ATTTGGCAGTCCGCTCTGATATATAGCAAGAACTTCAGACGATGCAAG 1796
|||||
400 yLeuGInGluAlaGluLeuArGAsPLeuValArGluAlaGluLysValT 417
|||||
1797 GTTACAGGAGGCTGAGTTACTGTGATCTTACGTACAGAGGCAGAGAGCTGT 1846
|||||
417 yTTrArGArGluThrGluGluGluLysGluGluInArGlySGluLysGlu 433
|||||
1847 ATTACAGAGGCGACACAGAGAGGAGAGACAGAGAAAGAGAAAGGAG 1896
|||||
434 ArGluGluArGluGluArGArGAsPArGArGInGluLysAsnLeuTh 450
|||||
1897 AGGAGAAAGAGGAGAGAAAGCGTGAATAGACGGCAAGAGAGATTTGAC 1946
|||||
450 rLysLLeuAlaAlaValAlaGluGlyLysSerSerArGluArGluA 467
|||||
1947 TAAAGATCTTGCCCGAGGTGTTGAAGGAGAGACACAGGAGAGAGAGA 1996
|||||
467 rAsPPhArGlyStLeArSerGlyProArGInSerGlyAsnLeuGly 483
|||||
1997 GAGATTTTAGGAAATTTAGTCAGGCCCCCTAGACAGTCAGAGAACTCGGGC 2046
|||||
484 AsnArGThrProLeuAsPlysAsPInCysAlaTyrCysLysGluLysG 500
|||||
2047 AATAGAGCCCCACCTCGAACAAGACACAGTGCCTATTGTAAAGAAAAAG 2096
|||||
500 yHisTrpAlaArGAsnCysProLysLysGlyAsnLysGlyProLysVal 517
|||||
2097 ACACCTGGCGCAAGAACTGCCCAAGAGGAGAAACAAGACGAGAGCTCC 2146
|||||
517 euAlaLeuGluGluAsPlysAsp 524
|||||
2147 TAGCTCTAGAAAGAGATTAAGAT 2169
|||||
seq_name: gb_pat:ARI30474

seq_documentation_block: 7333 bp DNA PAT 16-MAY-2001
LOCUS ARI30474
DEFINITION Sequence 2 from patent US 6190861.
ACCESSION ARI30474
VERSION ARI30474.1 GI:14118799
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 7333)
AUTHORS Fishman,J.A.
TITLE Molecular sequences of swine retroviruses method of using
JOURNAL Patent: US 6190861-A 2 20-FEB-2001;
FEATURES
Source Location/Qualifiers
1..7333
/organism="unknown"
BASE COUNT 1984 a 1788 c 1885 g 1676 t
ORIGIN
Alignment_scores:
Quality: 2758.00 Length: 525
Ratio: 5.263 Gaps: 1
Percent Similarity: 99.810 Percent Identity: 99.619
Alignment_block:
US-09-171-553B-4 x ARI30474 ..
Align seg 1/1 to: ARI30474 from: 1 to: 7333
1 MetcylcInThrValThrThrProLeuSerLeuThrLeuAspHsTrpTh 17
|||||
598 ATGGAGACGACAGTGAATACCCCTTAGTTGACTCTCGACCAATTGGAC 647
|||||
17 rGluValArgSerArgAlaHisAsnLeuSerValGlnValLysLysGly 34
|||||

```



```

648 TGAAGTTAGATCCAGGCGCTCAATTTGTGACGTTACAGGTTAAGAGGAC 697
34 rOTrGlnThrPheCysAlaSerGluTrpProThrPheAspValGlyTrp 50
698 CTTGGCACAACCTTCTGTGCTCTGATGCGCACATTCGATGATGTTGAGAGG 747
51 ProSerGluGlyThrPheAsnSerGluIleLeuAlaValLysAlaI 67
748 CCATAGAGGGGACCTTTAATTCTGAAATTTATCCGGCTGTTAAGGCAT 797
67 e1IlePheGlnThrGlyProGlySerHisProAspGlnGluProTyrIle 84
798 CATTTTTCAGACTGACCCGCTCATCTGATCAGAGAGCCCATATATCC 847
84 euThrTrpGlnAspLeuAlaGluAspProProProProTyrValLysProThr 100
848 TTACGCTGCACAACTTTGGCAGAAAGATCCTCGCCCATGGGTTAAACACAGC 897
101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGly 117
898 CTTAAATTAACCAAGAAAGCCAGGTCCTCCGAAATCCTGGCTTGGAGAGAA 947
117 sAsnLysHisSerAlaGluLysValGluProSerSerSerTyrLeuProA 134
948 AATCAACAACACTGCGCCGAAAGATGAGCCCTC.TCCTGATATCTACCCC 996
134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150
997 GAGATCGAGGAGCCGCCGACTTGGCCGAGAACCCCAACTGTTCCCCACAC 1046
151 ProLeuSerSerThrGlyCysCysGluGlyThr.SerAlaProProGly 167
1047 CCTTATTCACAGACAGGCTGCTGTAGAGGAGACCTCTGCCCTCTCGAG 1096
167 lApProValGluGlyProAlaAlaGlyThrArgSerArgArgGlyAla 183
1097 CTCGGTGGTGGAGGAGACTGCTGCCGGAGCTCGAGCCGAGAGAGGCC 1146
184 ThrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrG 200
1147 ACCCGGAGGAGACAGACGAGATCCGATATTACCGCTGCGACCTATGG 1196
200 yProProMetProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPhe 217
1197 CCTCCCATAGCCAGGGGCGCAATTCAGGCCCTCCAGATTTGGCCCTTT 1246
217 erSerAlaAspLeuTyrAsnTrpLysThrAsnHisProProPheSerG 233
1247 CTTTGTGACAGATCTTATATGTGAAACATTAACATCCCTTTCTCGAG 1296
234 AspProGlnArgLeuThrGlyLeuValGluSerLeuMetPheSerHisG 250
1297 GATCCCAAGCCCTCAGCGGGGTTGGAGTCCCTTATGTTCTCTCACCA 1346
250 nProThrTrpAspAspCysGlnGlnLeuLeuGlnThrLeuPheThrTrp 267
1347 GCCTACTTGGAGTATGTCACACAGCTGTCAGACACTCTTCACAAACCG 1396
267 lGluArgGluArgGlyLeuLeuGluAlaArgLysAsnValProGlyAla 283
1397 AGAGAGGAGAGAGAAATTCGTTAGAGGCTTGAATAAAATGTTCTCTGG 1446
284 AspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProle 300
1447 GACGGGCGACCCAGAGAGTTGCAAAATGAGATTTGACATGGATTTCCTT 1496
300 uThrArgProGlyTyrAspPyrAsnThrAlaGluGlyArgGluSerLeu 317
1497 GACTCGCCCGGTTGGACTACACACGAGCTGAAGGTAGGAGAGACTTGA 1546
317 ystLeuTyrArgGlnAlaLeuValAlaGlyLeuArgGlyLysSerArg 333
1547 AATTTATCGCCAGGCTGTGGTGGCGGTCTCCGGGGCGCTCAAGACGG 1596
334 ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGluPr 350
1597 CCCACTATTATTTGGCTAAGGTAAGAGAGTATCATCAGGAGCCGACGAC 1646
350 oProSerValPheLeuGluArgLeuMetGluAlaPheArgArgPheThr 367
1647 TCCCTCGGATTCTTTCAGAGGCTCATGGAAGCTTCAGGCGGTTCACCC 1696
367 rOpAspProThrThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 383
1697 CTTTGTATCTTACCTCAGAGGCCAGAAAGCTTCAGTGGCCCTGCGCTTC 1746
384 lIleGlyInSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGlu 400
1747 ATTGGCAGTCGCGCTTCGATATCAGAGAAACTTCAGAGACTGGAGAG 1796
400 yLeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGlyValT 417
1797 GTTACAGAGAGCTGATGATTACGTATGATGAGAGAGCCAGAGAGGCTGT 1846
417 yTyrArgArgGluThrGluGluGluGluArgGluArgGlyGlyGlu 433
1847 ATTACAGAGGAGAGACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1896
434 ArgGluGluArgGluArgGluArgGluArgGluGluGluGluGluGlu 450
1897 AGAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1946
450 rLysIleLeuAlaAlaValAlaGlyLysSerSerArgGluArgGlu 467
1947 TAAGATCTTGGCCGCTGATGTTGAGGAGAGAGAGAGAGAGAGAGAGAG 1996
467 rGAspPheArgGlyIleArgSerGlyProArgGlnSerGlyAsnLeuGly 483
1997 GAGATTTTGAAGAAATTAAGTCAGGCCCTTACAGAGTGAAGGAACTGG 2046
484 AsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluYsg 500
2047 AATGAGACCCACTCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2096
500 yHisTrpAlaArgAsnCysProLysGlyAsnLysGlyProLysVal 517
2097 ACACCTGGGCAAGGAACTGCCCAAGAGAGAGAGAGAGAGAGAGAGAG 2146
517 euAlaLeuGluGluAspLysasp 524
2147 TAGCTTAGAGAGAGATTAAGAT 2169
seq_name: gb_v1: PEN133818
seq_documentation block:
LOCUS PEN133818 8750 bp DNA VRL 14-APR-2000
DEFINITION Porcine endogenous retrovirus type C proviral gag, pol and env
genes and LTR (class B, clone 43).
ACCESSION AJ133818
VERSION AJ133818.1 GI:6688949
KEYWORDS env gene; env protein; gag gene; gag protein; pol gene; pol
protein.
SOURCE porcine endogenous retrovirus.
ORGANISM Porcine endogenous retrovirus
retroviruses; Retrod viruses; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
1 (bases 1 to 8750)
Czaderna,F., Fischer,N., Boller,K., Kurth,R. and Tonjes,R.R.
Establishment and characterization of molecular clones of porcine
endogenous retroviruses replicating on human cells
J. Virol. 74 (9), 4028-4038 (2000)
JOURNAL MEDLINE 20219394
REFERENCE 2 (bases 1 to 8750)
AUTHORS Toenjes,R.R.
TITLE Direct Submission
JOURNAL Submitted (04-MAY-1999) Toenjes R.R., Medical Biotechnology,

```

Paul-Ehrlich-Institut, Paul-Ehrlich-Strasse 51-59, Hessen, D-63225
Langen, GERMANY

FEATURES

Location/Qualifiers
source

/organism="porcine endogenous retrovirus"

/strain="type C"

/db_xref="taxon:61673"

/clone="43"

1..629

467

/note="cap site"

632..649

/note="gly4"

1064..2635

/gene="gag"

1064..2635

/gene="gag"

/function="retroviral capsid"

/product="Gag protein"

/note="viral polyprotein"

/protein_id="CAB65341.1"

/db_xref="GI:6688950"

/translation="MGCTVTPLSLIDHNTYKRSRAHNLVSVKKKPMOTFCASEMP

TFDGPSTGSEINSEILLAVKAIIFQTPSSHDDEYILITWDLADEPPWPKMIM

KPRKPGRIALGEKNKSAEKVEPSRIPELEPPPEPPPPPPAGAGAG

PSAPPGAVVEGPAAGRSRCAATPEKTDIELILPTRYPPMGLOLOWPSS

ADLYNMTNHPPESEDFORLGLVESLMSHPTMDCCOILLOTFTTERRELLIA

RKNVPGADGRPTOLONEIDMGFPLTRGMVNTABRESKITYROLVAGLRASRP

TNLAKEVEMGNEPSPVLEKMEFRFPPTSEKOKSAVALATGSDALDR

KTIQRLGLOEALRLVREAEVYTRKTEEEKERKEEREREKDRROKEMIT

KIIAAVEGKSSREEREDERRIRSGRQSONLNTPLDKDCAYCEKDHMAKPK

KGNKPGVIALBEDKGRGSDPLPERVALKEVGQVEELVTVGSHSLTLOPKL

KEKSWMGATGORYWTRRTVDLAVGRVTSFLVTPCPVPVILGRDLITKMGAI

SFGDREPVSNKPLITVLTLOLDEYRLSPQVKDOOIQSLDEPPQAMATYAGM

LAKOPOVYOLKASATPVSVROYLSREAREBTWVHVOIGILVYQSWNTPL

LPVRFETNDYRPVODIRENKRVDIHPVNPVNLALPERMWTYVLDLKAFF

CLRLHPTISPLFAFEMRDPGTGRTGOLTWTRLPQGNKSPITFDEALHDLNFRIOH

POVTLQYVDLLAGATKODCEGKALILELSDGYASAKAOICREVTYLGYS

LRGGQWMLTARKKTVQIAPPTAKOVBREIETAGFCRLMIPGFAITLAAPLYLKE

KGEENAPHEOKTFDAIKKALISAPALAPDYKPTLYVDEPKYAKARVLITOTGPM

RRPVATISKLDIVASGWPCLKAIAVAALVKKDADKLIGONTVITAHALENTVRO

PPRRMTNMTIYOSLLITERVTPAPALNATLIPETDEPVHDCOHLIEGT

VRDLDTDILTEGLVLTWFDSSVVEGRKMAAAYVDGTHIMASLDEGSAQAE

LEALTLPRKLAIECKSINIYDSRYAFATAHVAIKQGLTLAGREIKNEEILSL

ROYLTEDMOEIKIKIDOFSETPGCTYSYKTELIPKBELEVOOIHRTHTGTHLO

OLVRTSPYHVLRLPGVADSVKHCYPCQIYANPNSVPPKRLGSHPGAHWEVDTE
VPRAKGKNTILVVDTFSEWVAYPTKSTVAKKILEEIPFPFGIPKYGSDNG
PAFVAVOSGLAKILIGIDMKLHCAYRPOSSGOVERNAKIKLETLKTLTETINMIA
LLPFLVFRVNTPGQFLRPELLELGGPPLVEIASVHADVLISPLSRKALVEA
ROKAMKOLREAYSGEGDLQVPHRFQVDSVYBRHAGLLETRMGKSPYVILLITPPAY
KVEGISTWIAHVHLAAPPDGGWRAKTEENLAKTLHLHYVSNNSNGOMHPLISW
RHLPTGGEKRLRITLSTASTAMETLITITTOASAKRLIDSNRPLRSLTWLITDP
DTGVYTNSTRGVARCTWPELHFCRLINPAVKSTPPMLVSYKFCGCEKREKCY
GGSEESFCRMSVCVNSDNDKMPISLQDRVAFSVNSPGYKAYKLDKDGSCSPD
LDYLKISPEKROENIOKIMSGWIVFYKGGAGSTLILRIEIGTGPVAVG
PDKVLAEQGPALPEPHNLVPOLITSLRDLITPOGNSGTGLIIPNTRNSGPAVKT
GDLFLIOGAFOAINSTDPAATSSGWCISGSPYBEMAKEGFNITKEKNCTM
GSRNKTITLEVSGKCTICKAPPSHOHCISYVYEQASENOLVAPGINRWACHTGL
TPCVSSVFNOSKDFCVMOIVPVYHYHEEVVLEDEYDTRNPRAREVSLTAVMG
LGTAVGVGTGTAALITGPOLKEKGLGEHAHATBDRLALEESVLSLSEVL
ONRGLDLFLREGGICALKECCFYVDHSAITDSKALERRRERREDEOGM
FKGWRSPMTTLLSALTGLPLVLLLTITVGPCLINNFVATRRERVAVOIMVLRQ
YQGLTSQGETDI"

2639..6223

/gene="pol"

/function="retroviral protease and polymerase"

/product="pol protein"

2639..6223

/gene="pol"

mat_peptide 6099..8069
/gene="env"
/function="retroviral envelope"
/product="Env protein"

gene 6099..8069
/gene="env"

LTR 8122..8750
polyA_signal 8643..8648

BASE COUNT 2288 a 2214 c 2266 g 1962 t

ORIGIN

alignment_scores:
Quality: 2755.00 Length: 525
Ratio: 5.258 Gaps: 1
Percent Similarity: 99.810 Percent Identity: 99.619

alignment_block:
US-09-171-553B-4 x PEN133818 ..

Align seg 1/1 to: PEN133818 from: 1 to: 8750

1 MetGLyGlnThrValIThrThrProLeuSerLeuThrLeuAspHisTrpTh 17
|||||
1064 ATGGGACAGACAGTACGACACCCCTTATGTTGACTCTCCACCATTTGGAC 1113
17 rGLValArgSerArgLanHisAsnLeuSerValGlnValLysLysGlyP 34
|||||
1114 TGAAGTTAGATCCAGGCGCTCATATTGTGTAGTTCAAGTTAAGAGGAGC 1163
34 rOTrPGlnThrPheCysAlaSerGlnTrpProThrPheAspValGlyTrp 50
|||||
1164 CTGGCAGACTTCTGTGCTTGAATGCGCAACATTCATGTTGATGG 1213
51 ProSerGlnGlyThrPheAsnSerGlnIleIleLeuAlaValLysAlaI 67
|||||
1214 CCATCAGAGGGGACCTTAAATTCGAAATTAATCTGCTGCTGTTAAGCAAT 1263
67 eIlePheGlnThrGlyProGlySerHisProAspGlnGluProTyrIleL 84
|||||
1264 CATTTTTCAGACAGTGCACCCAGCTCTCATCTGATCAGAGCCCTTATATCC 1313
84 eutHrTrpGlnAspLeuAlaGluAspProProThrProTyrValLysProTrp 100
|||||
1314 TTACGTGGCAAGACTTGGCAGAAATCTCTCCGATGGGTTAAACCATG 1363
101 LeuAsnLysProArGlySerProGlyProArGlyIleLeuAlaLeuGlyL 117
|||||
1364 CTAAATTAACCAAGAAAGCCAGGCTCCCGCATCTGCTTGGCAGAGAA 1413
117 sAsnLysHisSerAlaGlnLysValGlnProSerSerSerTyrLeuProA 134
|||||
1414 AACCAAAACATCGCGCGGAAAAAATCCAGCCCTC.TCCTGTATCTACCCC 1462
134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150
|||||
1463 GAGATCGAGAGCGCGGCACTTGGCGGAACCCAAACGCTGTCCTCCACAC 1512
1463 GAGATCGAGAGCGCGGCACTTGGCGGAACCCAAACGCTGTCCTCCACAC 1512
151 ProLeuSerSerThrGlyCysGlyGlyThr.SerAlaProProGlyA 167
|||||
1513 CCTTATTCACGACAGAGGTCGTGTGAGGGGACCCTCTGCTCCGAGAG 1562
167 laProValValGlnGlyProAlaAlaGlyThrArgSerArgArgGlyAla 183
|||||
1563 CTCGGGTGTGGAGGAGCTGCTGCGGAGACTCGAGCGGAGAGAGCGCC 1612
184 ThrProGluArgThrAspGlnIleAlaIleLeuProLeuArgThrTyrG 200
|||||
1613 ACCCGGAGCGGACAGACAGATCGGATTAACCGCTGCGGACCATG 1662
200 yProProMetProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPhe 217
|||||
1663 CCTCCCATGCGCGGGGCGCAATTGACGCCCTCCAGATATGCGCCCTTTT 1712

217 eSerAlaAspLeuTyrAsnTrpAsnHisProProPheSerGlu 233
 |||||
 1713 CTTCTGAGATCTATATGAAACTAACCATCCCTTCTCGAG 1762
 234 AspProGlnArgLeuThrGlyLeuValGluSerLeuMetPheSerHisG1 250
 |||||
 1763 GATCCCAAGCGCTCAGCGGGTGGTGGAGTCCCTTATGTTCTCTCACCA 1812
 250 nProThrTrpAspPheCysGlnGlnLeuLeuGlnThrLeuPheThrhrG 267
 |||||
 1813 GCCTACTGCGATGATGTGCAACAGCTCTGAGAGACTCTTCACACCG 1862
 267 LgluArgGluArgIleLeuLeuGluAlaArgLysAsnValProGlyAla 283
 |||||
 1863 AGGACCGAGAGAAATCTGTAGAGGCTAGAAAAAAGTTCCCTGGGCC 1912
 284 AspGlyArgProThrGlnLeuGlnAsnGluLeuAspMetGlyPheProle 300
 |||||
 1913 GAGGGGCGACCGACAGTTCGCAAAATGATGACATGGATTTCCTT 1962
 300 uThrArgProGlyTrpAspTyrAsnThrAlaGluGlyArgGluSerLeu 317
 |||||
 1963 GACTGCCCCGGTGGAGCTACACAGCGCTGAGGTAGGAGGAGACTTGA 2012
 317 ySileYtrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArg 333
 |||||
 2013 AATCTATGCGCAGGCTGTGGGGGGGTCTCCGGGGCGCTCAAGAGCGG 2062
 334 ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGluPr 350
 |||||
 2063 CCCACTAATTTGGCTAGGTAGAGAGGTGATGACGAGGAGCCACAGAAC 2112
 350 oProSerValPheLeuGluArgLeuMetGluAlaPheArgPheThrP 367
 |||||
 2113 TCCCTCGATATTCTTGGAGAGGCTCATGAGACCTTCAGCGGTTCCACC 2162
 367 roPheAspProThrSerGluAlaGlnLysValSerValAlaLeuAlaPhe 383
 |||||
 2163 CTTTATCTCTACCTCGAGAGCCCGCAAGAGCTCACTGGCTTC 2212
 384 IleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGlu 400
 |||||
 2213 ATTGGGACAGTGGCTCTGATATCAGAAAGAACTTCAGAGCTGGAAG 2262
 400 yLeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluLysVal 417
 |||||
 2263 GTTACAGGAGCTGAGCTTACGTATCTAGTACAGAGGACAGAAAGTGT 2312
 417 yTrpArgArgGluThrGluGluGluGluGlnArgLysGluLysGlu 433
 |||||
 2313 ATTACAGAGGAGGACAGAAAGAGACAGAAAGCAAGAAAGAAAGAGAG 2362
 434 ArgGluGluArgGluGluArgArgAspArgArgGlnGluLysAsnLeuTh 450
 |||||
 2363 AGAGAGAGAAAGGAGGAGAAAGCTGATAGAGCGCAAGAGAAATTTGAC 2412
 450 rLysIleLeuAlaAlaValAlaGluLysSerSerArgGluArgGluAla 467
 |||||
 2413 TAGATATGCGCCGACGTCTTAAAGGGAAGACAGCAGGAGAGAGAGA 2462
 467 rGAspPheArgLysIleArgSerGlyProArgGlnSerGlyAsnLeuGly 483
 |||||
 2463 GAGATTTTAGGAAATAGGTACAGGCCCTAGACAGTACAGGAGAACTGGGC 2512
 484 AsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysG1 500
 |||||
 2513 AATTAGACCCCTACTGCAAGAGACCAAGTGTGCGATTGTAAAGAAAAAGG 2562
 500 yHisTrpAlaArgAsnCysProLysLysGlyAsnLysGlyProLysVal 517
 |||||
 2563 ACACCTGGGCAAGAGACTGCCCAAGAGAGGAAACAAAGAGCCAGAGGTCC 2612

517 euAlaLeuGluGluAspLysAsp 524
 |||||
 2613 TAGCTCTAGAGAGAGATTAAGAT 2635

seq_name: gb_v1: PEN293657

seq_documentation_block: 8763 bp DNA VRL 30-MAY-2001

LOCUS PEN293657 Porcine endogenous type C retrovirus proviral gag gene, pol gene

DEFINITION and env gene, class B, clone 213.

ACCESSION AJ293657 GI:14275841

VERSION env gene; envelope; gag gene; group specific antigen; pol gene;

KEYWORDS polymerase; polypotein.

SOURCE porcine endogenous type C retrovirus.

ORGANISM Viruses; Retroid viruses; Retroviridae; Gammaretrovirus.

REFERENCE 1 (bases 1 to 8763)

AUTHORS Krach, U., Fischer, N., Czauder, F. and Tonjes, R. R.

TITLE Comparison of replication-competent molecular clones of porcine

JOURNAL endogenous retrovirus class a and class b derived from pig and

human cells

MEDLINE J. Virol. 75 (12), 5465-5472 (2001)

REFERENCE 2 (bases 1 to 8763)

AUTHORS Tonjes, R. R.

TITLE Direct Substitution

JOURNAL Submitted (31-JUL-2000) Tonjes R. R., Medical Biotechnology,

Paul-Ehrlich-Institut, Paul-Ehrlich-Str. 51-95, Langen, 63225,

GERMANY

FEATURES

source

Location/Qualifiers

1..8763

/organism="porcine endogenous type C retrovirus"

/proviral

/specific_host="Sus scrofa"

/db_xref="taxon:70540"

/clone="213"

/cell_line="PK15"

/note="class B"

1..630

467

/note="Cap site"

633..650

1077..2651

/gene="gag"

1077..2651

/gene="gag"

/codon_start=1

/product="group specific antigen"

/protein_id="CAC39625.1"

/db_xref="GI:14275842"

/translation="MQQVYTPPLSLTLDHMTVEYRSRAHNLVQKKGPWOTFCASEMP

TFDVGMPSEGTENSEILIAVKAIIPOTGSHSDDEPYILTWODLAEPPPWKFWLN

KPRKPGRIALIGENKNSHAEKVPSPGILYELIEPPWPEQPVPPYPAGAVRG

PSAPPCGAVVEGPAAGTRSRGATPRTEIALPLRTYGPMPGGLOLPLOYWPS

ADLYNKNTNRPSPEDPORLTVGIVESIMSHOPTMDCOOLLOTLETEREILIEA

RKNVPGADGRPTDLDNEDMGFPILTRPMQDYNFAEERESIKYIRQLAVGRCASRP

TNLAKYREVMQGNPPSVFLERLMEAFREFTFEDTSEQKASVALATIGOSALDIR

KILQRLGLOEALRDLYRAEKVYRRETEERKEQREEREREERDRROEKULT

KTLAAVEGKSSRERDRFKIRSGPRSGNLGNRTPLDKDCAAYCKEKGHWARNCPK

KGNGKPVLALEDKDK"

2652..6239

/gene="pol"

<2652..6239

/gene="pol"

/note="complete pol ORF, gag-pol is expressed as a

polyprotein"

/citation=[1]

/codon_start=1

/product="polymerase"

/protein_id="CAC39626.1"

/db_xref="GI:14275843"

/translation="GRGSDPLPEPRVTLKVEQGPVEFLVDTGAEHSVLLQPLGLIKE

[illegible]

1277	CATTTCACGACTGGACCCGGCTCTCATCTGTATCGAGAGCCCTATATCC	1326
84	euthrTPGInAspLeuAlaGluAspProProtrPValysProtrP	100
1327	TTACGTGGCAAGATTGTGGCAAGAAACCTCCGGCAATGGGTAAACCATGG	1376
101	LeuAsnLysProArgLysProGlyProArgLLeuAlaLeuGlyGly	117
1377	CTAAATTAACCAAGAAAGCCAGGTCGCCGATCTGGCTTGGAGAGAGA	1426
117	sAsnLysHisSerAlaGluLysValGluProSerSerSerTLeuProA	134
1427	AAACAACACCTGGCGCAAAAAATCCAGACCTCTCTCTGG. TATCTTAACCC	1475
134	rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProthr	150
1476	GAGATTCAGAGAGCGCCGACTTGTGGCCGGAACCCCAACTGTGCCCCACC	1529
151	ProLeuSerSerThrGlyCysCysGluGlyThr. SerAlaProGlyA	167
1526	CCCTTATCCAGACAGGGTGTGTGAGGGAGCCATCTGCCCTCTCTGGAG	1575
167	laProValAlaGluGlyProAlaAlaGlyThrArgSerArgArglyAla	183
1576	CTCCGGGGTGGAGGAGACTGTCTGGCGGACACGGAGCGGAGAGCGCC	1625
184	ThrProGluArgThrAspGlyTLeuAlaAlaLeuProLeuArgThrTyrG	200
1626	ACCCGGAGGGAGACAGCAAGATCCGATATTACCGCTGGCCACTTATGG	1675
200	YProProMetPProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProhes	217
1576	CCCTCCCATGCGAGGGGGCAATTGCAGCCCTCCAGATATGGCCCTTTT	1725
217	eSerAlaAspLeuTyrAsnTrpLysTrpAsnHisProProPheSerGlu	233
1726	CTTCTGCAGATCTCTATATTGGAAAACTAAACATCCCTCTTCTGGAG	1775
234	AspProGlnArgLeuThrGlyLeuValGluSerLeuMetPheSerHisG	250
1776	GATGCCCAACGGCTACAGGGGTGTGGAGTCCCTTATGTTCTCTACCA	1825
250	nProThrTrpAspAspCysGlnGlnLeuGlnGlnThrLeuPheThrThrG	267
1826	GCTTACTTGGATGTATGTCAACACTCTCTCAACAACCTTCAACAACG	1875
267	LuGlnArgGlnArgLLeuLeuGlnGluAlaArgLysAsnValProGlyAla	283
1876	AGGAGCGAAGAGAAATTGCTTAAGAGCTTAACAAAAATGTTCTCGGGGCC	1925
284	AspGlyArgProThrGlnLeuGlnAsnGluLeuAspMetGlyPheProle	300
1926	GAGGGCGACCCACCCASCTTCCAAATATGATATGACATGGGATTTCCCTT	1975
300	uThrArgProGlyTyrTrpAspTyrAsnThrAlaGluArgLysSerLeuL	317
1976	GACTCGCCCCGGTGTGGACTCAACACGGCTGAAGAGTGAAGAGCACTGA	2025
317	ysLeuTyrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArg	333
2026	AAATCTATACGACAGCTCTGTGGCGGAGTCCCGGGCGGCTCAACAGG	2075
334	ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGluPr	350
2076	CCCACTAATTGGCTACAGTAAGAGAGGTATACAGGAGCCGAACCAACCC	2125
350	oProSerValPheLeuGlnArgLeuMetGluAlaAlaPheArgArgPheThrP	367
2126	TCCCTCGGATTTCTTGAGAGGCTCATGAAAGCCTTCAGGCGGTTACCC	2175
367	roPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe	383

```

2176 CTTTGATCTACCTCAGAGGCCGAGAAAGCCTCAGTGGCCCTGGCCTC 2225
384 IleglyglnSerAlaIleuAspIleArgLysLysLeuGlnArgLeuGlu 400
2226 ATTGGGCGATGCGCTCTGATATCAGAAAGAACTTCAGAGCTGAGAGG 2275
400 yLeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluLysVal 417
2276 GTTACAGGAGGCTGAGTTACGATCTAGTGAGAGGAGCAGAGAGGTGT 2325
417 yTyrArgArgGluThrGluGluGluGluGluGluGluGluGluGlu 433
2326 ATTACAGAAAGGAGACAGAGAGGAGAGAAAGAAACAGAAAGAGAG 2375
434 ArgGluGluArgGluGluGluGluGluGluGluGluGluGluGlu 450
2376 AGAGAAAGAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2425
450 yLysIleLeuAlaIleValIleGluGluLysSerSerArgGluArgGlu 467
2426 TTAGATCTTGCGCGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAG 2475
467 rGAspPheArgLysIleArgSerGlyProArgGlnSerGlyAsnLeuGly 483
2476 GAGATTTTGAAGAAATTAGTCAGGCCCTAGACAGTCAGGGAACCTGGGC 2525
484 AsnArgThrProLeuAspLysAspGlnCysAlaIleTyrCysLysGluLysGlu 500
2526 AATAGAGCCCACTCGACACAGAGAGAGAGAGAGAGAGAGAGAGAG 2575
500 yHisThrAlaArgAsnCysProLysLysGlnLysAsnLysGlyProLysVal 517
2576 ACACCTGGCAGAGAACTGCCCCAGAGAGAGAGAGAGAGAGAGAGAG 2625
517 euAlaLeuGluGluAspLysAsp 524
2626 TACCTAGAGAGAGATTAAGAT 2648

seq_name: gb_v1:PEN133816

seq_documentation_block:
LOCUS      PEN133816      8918 bp      DNA      07-DEC-2000
DEFINITION Porcine endogenous retrovirus type C proviral gag, pol and env
genes and LTR (class B, clone 33).
ACCESSION AJ133816
VERSION    AJ133816.1 GI:6688945
KEYWORDS   env gene; env protein; gag gene; gag protein; pol gene; pol
           protein.
SOURCE      porcine endogenous retrovirus.
ORGANISM   Porcine endogenous retrovirus
            Viruses; Retrovirdae; Mammalian type C
            retroviruses; 1-Mammalian type C virus group.
REFERENCE   1 (bases 1 to 8918)
AUTHORS     Czadudena,F., Fischer,N., Boller,K., Kurth,R. and Tonjes,R.R.
TITLE       Establishment and characterization of molecular clones of porcine
            Establishment and characterization of molecular clones of porcine
            Establishment and characterization of molecular clones of porcine
            Establishment and characterization of molecular clones of porcine
JOURNAL     J. Virol. 74 (9), 4028-4038 (2000)
MEDLINE     20219394
REFERENCE   2 (bases 1 to 8918)
AUTHORS     Tonjes,R.R.
TITLE       Direct Submission
JOURNAL     Submitted (04-MAY-1999) Tonjes R.R., Medical Biotechnology,
            Paul-Ehrlich-Institut, Paul-Ehrlich-Strasse 51-59, Hessen, D-63225
            Langen, GERMANY

FEATURES
            source
            1..8918
            /organism="porcine endogenous retrovirus"
            /proviral
            /strain="type C"
            /db_xref="taxon:61673"
            /clone="33"
            /note="grown in PK15 (ATCC No. CCL-33) cell line"

LTR
1..707
BASE COUNT 2327 a 2245 c 2334 g 2012 t

```

ORIGIN

alignment_scores: Length: 525
 Quality: 2751.00
 Ratio: 5.250
 Gaps: 1
 Percent Similarity: 99.810 Percent Identity: 99.429

alignment_block:
 US-09-171-553b-4 x PEN133816 ..

Align seg 1/1 to: PEN133816 from: 1 to: 8918

```

1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisTrpH 17
1154 ATGGGACAGACAGTGCAGACCCCTTACTTAACTCGACCAATTGGAC 1203
17 rGluValArgSerArgAlaHisAsnLeuSerValGlnValLysLysGlyP 34
1204 TGAAGTTAATCCAGGGCTCATTAATTGTCAAGTTCAGGTAAAGAGGAC 1253
34 rOTrPGLnThrPheCysAlaSerGluTrpProThrPheAspValGlyTrp 50
1254 CTGGCAGACTTCTGTGCTCTGATGCAATGCCAATTCGATGTGGATGG 1303
51 ProSerGluGlyThrPheAsnSerGluIleIleLeuAlaValLysAla11 67
1304 CCATCGAGAGGGACCTTAAATCTGAATAATATCTGGCTGTTAAGCAAT 1353
67 eIllePheGlnThrGlyProGlySerHisProAspGlnGluProtyrIleL 84
1354 CATTTTTCAGACTGCAGCCAGCTCTCATCTCATCGACGACCCCTATATCC 1403
84 eUrThrTrpGlnAspLeuAlaGluAspProProProTrpValLysProTrp 100
1404 TTACGGCGCAGAGACTTGGCAGAAATCCCGCCCATGGGTTAAACCATGG 1453
101 LeuAsnLysProArgLysProGlyProArgGlyIleLeuAlaLeuGlyLulY 117
1454 CTAAATAAACCAAGAAAGCCAGGTCGCCGAATCCGCTCTGGAGAGAA 1503
117 sAsnLysHisSerAlaGluLysValGluProSerSerSerTyryleuProA 134
1504 AAACAACACTGCGCGCAAAAAGTGGAGCCCTC.TCCTCGATATCTACCCC 1552
134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150
1553 GAGATCGAGGAGCCCGCACTTGGCCGGAACCCCAACCTGTTCCCCACC 1602
151 ProLeuSerSerThrGlyCysCysGluGlyThr.SerAlaProProGlyA 167
1603 CCTTATCCAGCACAGGGTCTGTGAAGGGACCCCTCTGCCCTCTGGAG 1652
167 lAProValAlaGluGlyProAlaAlaGlyThrArgSerArgGlyAla 183
1653 CTCGGGTGTGAGGAGACCTGCTGCGGGGACTCGAGAGCCGGAAGGGGCC 1702
184 ThrProGluArgThrAspGluIleAlaAlaLeuProLeuArgThrTyrgl 200
1703 ACCCGAGGCGGACAGAGAGATGCGATATTACCGCTGCGCACACTATGG 1752
200 yProProMetProGlyGlyGlnLeuGlnProLeuGlnTyTrpProHes 217
1753 CCCTCCCATGGCGGGGGGCCAATTGCAAGCCCTCCAGATATTGGGCCCTTTT 1802
217 eTSerAlaAspLeuTyAsnTrpLysThrAsnHisProProHesSerGlu 233
1803 CTCTCGAGATCTCTATAATTGAAACTAACAATCCCTCTTCTGAGAG 1852
234 AspProGlnArgLeuThrGlyLeuValGlnSerLeuMetHesSerHisgl 250
1853 GATCCCAACGCGCTCACGGGGTGTGTGAGTCCCTATGTCTCTACCA 1902

```

```

250 nProThrTrpAspAspCysGlnGlnLeuGlnThrLeuPheThrTrng 267
1903 GCCTACTTGGGATGATGTGTCAACAGCTCTGACAGACACTTTCACACCG 1952
267 lUglUArgGluArgIleLeuLeuGlnAlaArgLysAsnValProGlyAla 283
1953 AGGACGAGAGAGAAATCTGTAGAGGCTAGAAAAAATGTTCTGTGGGCC 2002
284 AspGlyArgProThrGlnLeuGlnAsnGluLeuLeuAspMetGlyPheProLe 300
2003 GACGGGCGACCCACGCACTTGCATAAATGAGATTGACATGAGGATTTCCCTT 2052
300 uThzArgProGlyTrpAspTyraSnThrAlaGluIleValArgGluSerLeuL 317
2053 GACTGCGCCGGTGGGACTACCAACACGCGCTGAAGGTRAGGAGACTTGA 2102
317 ySlleTyArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArg 333
2103 AAATCTATCGCCAGGCTCTGTGGCGGTCTCCGGGGCGCCTCAAGACGG 2152
334 ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGluTr 350
2153 CCCCTAATTTGGCTAAAGTAAAGAGGTATGACAGGAGCGAAGCAAC 2202
350 oProSerValPheLeuGluArgLeuMetGluAlaPheArgArgPheThrP 367
2203 TCCCTCGGTATTCTTGAAGAGCTCATGGAAGCTTACGGCGGTTCACCC 2252
367 rOPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 383
2253 CTTTGAATCTTACCTCGGAGGCCGGAAGACCTCAGTGGCCCTGGCCTTC 2302
384 lLeuGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGlu 400
2303 ATTGGGAGTGGCTCTGATATCAGAAAGAACTTCAGAGCTGGAAGG 2352
400 yLeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluLysValTr 417
2353 GTTACAGAGAGCTGAGTTAGCTGATCTAGTGAAGAGCGAGAAAGTGT 2402
417 yTyrrArgArgGluThrGluGluGluLysGluGlnArgLysGluLysGlu 433
2403 ATTACAAAGAGGAGACAGAAAGAGAGAGAAAGCAAGAAAAAGAGAG 2452
434 ArgGluGluArgGluArgGluArgAspArgGluGlnLysAsnLeuH 450
2453 AGGAAGAAAGAGAGAAAGACCTGATAGACGCCAAGAGAAATTTGAC 2502
450 rLysIleLeuAlaAlaValAlaGluGlyLysSerSerArgGluArgGluA 467
2503 TAAAGATCTTGGCCGCACTGTGTGAAGGAAAGACAGACAGAGAGAGAGA 2552
467 rGAspPheArgLysIleArgSerGlyProArgGlnSerGlnLeuGly 483
2553 GAAATTTTGAAGAAATTAAGTCAAGCCCTAGACAGTCAAGGAACTGGGCC 2602
484 AsnArgThrProLeuAspLysAspGlnCysAlaTyCysLysGluLysgl 500
2603 AATAGAGCCCACTGACAAAGACCAAGCTGCTGATATTGAAGAAAAAG 2652
500 yHisTrpAlaArgAsnCysProLysLysglLysAsnLysglProLysValL 517
2653 ACACCTGAGCAAGGAACTGCCCAAGAGGAAACAAAGAGACGGAAGTCC 2702
517 euAlaLeuGluGluAspLysasp 524
2703 TAACTCTAGAGAGATTAAGAT 2725

```

seq_name: gb_v1:PERY17013

seq_documentation_block: 7808 bp DNA VRL 14-APR-2000

LOCUS PERY17013

DEFINITION Porcine endogenous retrovirus gag, pol, env genes.

ACCESSION Y17013
 VERSION Y17013.1 GI:3005985
 KEYWORDS env gene; gag gene; pol gene.
 SOURCE porcine endogenous retrovirus.
 ORGANISM Porcine endogenous retrovirus
 Viruses; Retroviridae; Mammalian type C
 retroviruses; 1-Mammalian type C virus group.
 1 (bases 1 to 7808)
 Czauderna, F., Fischer, N., Boller, K., Kurth, R. and Tonjes, R.R.
 Establishment and characterization of molecular clones of porcine
 endogenous retroviruses replicating on human cells
 J. Virol. 74 (9), 4028-4038 (2000)
 20219394
 2 (bases 1 to 7808)
 Tonjes, R.R.
 Direct Submission
 Submitted (30-MAR-1998) R.R. Tonjes, Paul-Ehrlich Institut,
 Paul-Ehrlich Str. 51-59, P.O. Box 1740, D-63207 Langen, FRG
 Location/Qualifiers
 1. 7808
 /organism="porcine endogenous retrovirus"
 /db_xref="taxon:61673"
 /tissue_type="kidney"
 /cell_line="PK15"
 /gene="gag"
 260..1834
 /gene="gag"
 260..1834
 /gene="gag"
 /codon_start=1
 /protein_id="CAA76581.1"
 /db_xref="GI:3005986"
 /db_xref="SPTREMBL:O73504"
 /translation="MGQVYTPISLTLDHMTVEVRSRAHNLVSVQKKGPMOTFCASEMP
 TFDVGPSEGTNSITLAVKALITPOTGSSHPDQSLITLQDLAEDPPPVKVMNLN
 KPRPGPRLILAGEKNKHSAREKVEPRILYEPPEPPEPPPPYPAQAVRG
 PSAPGAPVVEGPACTRSRGATPRTDEIALPLRTYPPGGQLQDLYPERS
 ADLYNKTNPRESDEPRLTGIVESTLME SHO PWDCCQLQTLFTAEERILIEA
 RNKVPAGDGPQLONEIDMGWPLTFLRWYNAEGRESKILYRQALVLAGASRP
 TNIAKRYVMGNGPPEPPEPPEPPEPPEPPEPPEPPEPPEPPEPPEPPEPPEP
 KTLQRLGLOEALRLVREAEKATYGRTEPEKREKEREKEREKEREKEREKEREK
 KILAAVEGSSKREKREKREKREKREKREKREKREKREKREKREKREKREKREK
 KGNKGPVLALEDKD"
 2120..4726
 /gene="pol"
 2120..4726
 /gene="pol"
 /codon_start=1
 /protein_id="CAA76582.1"
 /db_xref="GI:3005987"
 /db_xref="SPTREMBL:O73505"
 /translation="MGAQISFEQGRPEVSNKPLITLTLQDLDEYRLSPQVKKPDOD
 IQSMLDEFPQAMAEAGMGLAKQVPPQVQIOLKASATPVSVROYPLSRERAGIWPVO
 RLIQGTLVPOSPWNTPLIVRRPGTNDIRPVODLREYNKRVODIHPPVNPYILS
 ALPERMNVVLDLKDAPFLRLHPTSPDLAEWMDPGTGRGOLITWRLDGFKNS
 PTJFEDLHRDLANFRIOHPQVTLQYVDLLAGATGAPCTEGTALLIELSDLYR
 ASAKKQICREVTLYGYSLRGQRMUTLARKKTVQIAPPTAKQVREFLIAGACR
 LMIIGATLAPLYPLTKEKGEFSNAPEHOKADAIKALISAPALADVDKPTLY
 VDERKYGARVLIOTLGPMPRPVAYLSKTIAPVASCMPYCIKAAVALIVDADKLT
 LGQNTIVIAHLENTVROPDRMTNARWTHIQSLILTERVTFAPPALNPITLPE
 ETDEPVHDCQHLIETGVKDLIDPLIGEVLTWFTDQSSVYVGGKRNAGAAYVDG
 THTIMASLPECHTSAQAEMLALTOALRLAEGKSINITYDSRYAFATVHGAIVQR
 GLTISAREIKNEELISLEALHLEKRLAIHCPGHOKAKDLISGRNOMADVAAQA
 AOAENVLPIETPRAPRRROYLTLEDMOEIKIDQSEPECTCYISYGEKILPHEG
 LEVYQOHRITHTGKTKIOQLVRSYHVLRLGVADSVYKHCVPOLNANANSTRILP
 GKRIRGSHGPAHVEVPTTEVKAIPKTKYLLVVDVFSGWEAVPTKKEISIVAKYI
 WRKRFDELYLR"
 5294..7267
 /gene="env"
 5294..7267
 /gene="env"
 /codon_start=1
 /protein_id="CAA76583.1"
 /db_xref="GI:3005988"

/db_xref="SPTREMBL:O73506"
 /translation="MHPILSMRHLPTVGGEPKRLRLDLPASIAIMELTTLTPQASSK
 RLIDSSNHPRLSLTLILIDPDGVAVSTRGVAPRGTMPELFLRLINPAVKSTP
 NLIVSYGFYCCPGCTEKEXYQSGSESEFCRMSCTVNDQDKWPILODRVAFSVN
 SGPKRYKMLYKXKSCSPSDIDVLTISPEKQCEQENIOWKMSGIVFYVGGGA
 GSTLILRLLETGTPEVAVGPBVDVLAEGPPALEPHNLPPVPLTSLRDTIOPSN
 GTGLIPTNPVNSPGVAVKGTGRFLSLIGAVQAINSTPDATSSCWLISGPPY
 EGMAEKGFVNTKEHNOCTWGSRNKLITLVEGSKGCIKAPSHQHLICSTVYEQ
 ASENQVLPQGNRMWACNTGLTPCVSSSVFNSQKFCVMQIVPRVYHBEVVDLEY
 DYRNRPKREPVSLLIAMIGLGAIVSGVGTALITPOOLEKGLGELHAAMTEDLR
 ALESVSNLESLSLSEVIONRGIDILFLRGGICALLKECCCFYVHSGAIRS
 MSKLRLERRRRRRREADQGFECGRFNRSPMTLTLALIGPLVLLILLTVGECILN
 RFVAFVERVGAQVQIMVLDQYGGILSGGETDL"
 BASE COUNT 2077 a 1971 c 2047 g 1713 t
 ORIGIN
 alignment_scores:
 Quality: 2729.00 Length: 525
 Ratio: 5.228 Gaps: 1
 Percent Similarity: 99.429 Percent Identity: 98.857
 alignment_block:
 US-09-171-553b-4 x PERY17013 ..
 Align seg 1/1 to: PERY17013 from: 1 to: 7808
 1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisThrPth 17
 260 ATGGGACAGACAGTGCAGACCCCTTAGTTACTCTGCACCATTTGGAC 309
 17 rglValArgSerArgAlaHisAsnLeuSerValGlnValLysGlyP 34
 310 TGAAGTAGATCCAGGCGCATATTGTTCAGTTACGTTAAGAGGAC 359
 34 rotTrpGlnThrPheCysAlaSerGluTrpProThrPheAspValGlyTrp 50
 360 CTTGCGCAGCTTCTCTGCTCGCTGAAGGCAATTCGATTCGATGGATGG 409
 51 ProSerGluGlyThrPheAsnSerGluIleIleLeuAlaValLysAlaI 67
 410 CCATCAGAGGCGACCTTAATTTCTGAATATATCTGCTGTTAAGGCAAT 459
 67 eilePheGlnThrGlyProGlySerHisProAspGlnGluProTyrIle 84
 460 CATTTCAGACTGACCGACCTCTCATCTGATCAGAGACCCCTATATCC 509
 84 eutThrTrpGlnAspLeuAlaGluAspProProProTyrValLysProTyr 100
 510 TTACGTGCGCAAGACTTGGCAGAAAGATCTCCGCGCATGGTTAAACCATGG 559
 101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGly 117
 560 CTAAATTAACCAAGAAAGCCAGGTCGCCGATCTGCTTGGGGAGAA 609
 117 sAsnLysHisSerAlaGluLysValGluProSerSerSerTyrLeuPro 134
 610 AAACAACACCTCGCGCGAAAAAGTCGAGCCCTC.TCCTCGTATCTACCC 658
 134 rgaAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150
 659 GAGATCAGAGGCGCGGACTTGGCGGAACCCCAACCTGTTCCCCACC 708
 151 ProLeuSerSerThrGlyCysCysGlyGlyThr.SerAlaProProGly 167
 709 CCCTTATTCACAGACAGGCTGTGTGAGGAGACCTCTGCCCTCTGAG 758
 167 laProValValGlyGlyProAlaAlaGlyThrArgSerArgArgGlyAla 183
 759 CTCGGGTGTGTGAGGAGACCTGCTCGCGGATCGGACCGAGAGCGCC 808
 184 ThrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrG 200

809 ACCCCGAGCGGACAGACGATCGGATATACCGCTGGCCACTATGG 858
 200 YPRPROMETPROGLYGLYINLEUINPROLEUINTRYTPPROPHES 217
 859 CCTCCCATCCCGGGGGCCAAATTGACGCCCTCCAGTATGGCCCTTT 908
 217 exSerAlaAspLeuTYrAsnTrpLysThrAsnHisProProPheSerGlu 233
 909 CTCTGACAGATCTCTATTAATGGAAACTAACCATCCCTCTCTCGGAG 958
 234 AspProGlnArgLeuThrGlyLeuValGlnSerLeuMetPheSerHisG1 250
 959 GATCCCAACAGCCCTACGGGGTGGAGTCCCTATATGTCCTCACCAC 1008
 250 nProThrTPASPAPSPCYSGINGINLEUENGINHTRLEUPHERTHRhg 267
 1009 GCTTACTTGGAGATTGTCAACAGCTGCTGCAGACACTCTTCACAGCCG 1058
 267 LUGLUAArgGLUAArgLLeuLeuGluAlaArgLysAsnValProGluAla 283
 1059 AGAGAGAGAGAGAAATTCCTTAGAGCCTAGAAAAAATGTTCTGGGGCC 1108
 284 AspGlyArgProThrGlnLeuGlnAsnGluLeaspMetGlyPheProLe 300
 1109 GACGGGCGACCCACGACGTTGCAAAATGAGATTGACATGGGATTTCCCTT 1158
 300 uThrArgProGlyTrpAspTYrAsnThrAlaGlnGlyArgGlnSerLeuL 317
 1159 GACTCGCCTCGGTTGGACTACACACGCGCTGAAGAGTAGGAGAGCTTGA 1208
 317 ySILeTYrArgGLAlaLeuValaGlyLeuArgGlyAlaSerArgArg 333
 1209 AAATCTATCGCCAGGCTCTGGGGGGCTCCCGGGGGCCCTCAAGACGG 1258
 334 ProThrAsnLeuAlaLysValaArgGluValMetGlnGlyProAsnGluPr 350
 1259 CCCACTAATTGGCTAAGGTAAAGAGGTGATGCAGGACCGAAGCAAC 1308
 350 oProSerValPheLeuGluArgLeuMetGluAlaPheArgArgPhehrp 367
 1309 TCCCGGTATTTCTTGAGAGGCTCATGAGAGCTTCAGGGGGTTCCACC 1358
 367 roPheAspProThrSerGluAlaGlnLysAlaSerValaLeuAlaPhe 383
 1359 CTTTGTATCTTACCTCGAGAGCCAGAAAGCTCAGTGGCCCTGGCTTC 1408
 384 ILleGlyInSerAlaLeuAspLLeArgLysLysLeuGlnArgLeuGluG1 400
 1409 ATTGGGAGTGGCTGCTGATATCAGAAAGAACTTCAGAGACTCGAAGG 1458
 400 yLeuGlnGluAlaGluLeuArgAspLeuValaArgGluAlaGluLysValT 417
 1459 GTTACAGAGAGGCTGAGTTAGCTGATCTAGTGAAGAGGAGAGAGGCT 1508
 417 yTYrArgArgGluThrGluGluLysGluGlnArgLysGluLysGlu 433
 1509 ATTAACAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1558
 434 ArgGluGluArgGluGluArgArgAspArgArgGlnGluLysAsnLeuTh 450
 1559 AGAGAAAG 1608
 450 rLYsILeLeuAlaAlaValaValaGluLysSerSerArgGluArgGlu 467
 1609 TAGATCTTGGCCGAGTGGTTAAGAGAGAGAGAGAGAGAGAGAGAGAGA 1658
 467 rGAspPheArgLysLLeArgSerGlyProArgGlnSerGlyAsnLeuGly 483
 1659 GAGATTTTGAAGAAATTAAGTTCAGGCCCTAGACAGTCAAGAGAACCTGG 1708
 484 AsnArgThrProLeuAspLysAspGlnCysAlaTYrCysLysGluLysG1 500
 1709 AATAGAGCCCACTCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1758

500 yHISrPAlaArgAsnCysProLysGlyAsnLysGlyProLysValL 517
 1759 ACACGGGCGAAGAGAACTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1808
 517 euAlaLeuGluGluAspLysAsp 524
 1809 TAGCTCTAGAGAGAGATTAAGAT 1831

seq.name: gb_v1:PEN293656

seq_documentation_block:

LOCUS PEN293656 8918 bp DNA VRL 30-MAY-2001
 DEFINITION Porcine endogenous type C retrovirus proviral gag gene, pol gene
 and env gene, class A, clone 58.

ACCESSION AJ293656

VERSION AJ293656.1 GI:14275837

KEYWORDS env gene; envelope; gag gene; group specific antigen; pol gene;
 polymerase; polypeptide.

SOURCE Porcine endogenous type C retrovirus.

ORGANISM Viruses; Retrovirus; Retroviridae; Gammaretrovirus.

REFERENCE 1 (bases 1 to 8918)
 Krach, U., Fischer, N., Czauderna, F. and Tonjes, R.R.
 Comparison of replication-competent molecular clones of porcine
 endogenous retrovirus class a and class b derived from pig and
 human cells

J. Virol. 75 (12), 5465-5472 (2001)

JOURNAL MEDLINE 21256017

REFERENCE 2 (bases 1 to 8918)
 Tonjes, R.R.
 Direct Submission

AUTHORS Submitted (31-Jul-2000) Toenjes R.R., Medical Biotechnology,
 Paul-Ehrlich-Institut, Paul-Ehrlich-Str. 51-95, Langen, 63225,
 GERMANY

TITLE JOURNAL

FEATURES

source

Location/Qualifiers

1..8918

/organism="porcine endogenous type C retrovirus"

/proviral

/specific_host="Sus scrofa"

/db_xref="taxon:70540"

/clone="58"

/cell_line="PK15"

/note="class A"

1..702

LTR

misc_feature

467

/note="Cap site"

706..723

1153..2727

/gene="gag"

1153..2727

/gene="gag"

/codon_start=1

/product="group specific antigen"

/protein_id="C639616.1"

/db_xref="GI:14275838"

/translation="MGQVYTPPLSLFDHMTVEKSRANLSVQYKKGPMQTESEMP
 TFDVGPSECTFSESLIAVKATIPOTGSHDPOPYLTQDLADEPFPVYKFWLN
 KPRPGPRLIALGEMKHSAEKPSPHIYELPEPPAPPEQSVPPPAQAGAR
 PSAAPGAPAVEGAACTRSRGATPERDEIATLPLRYGPPIPGQLOPLDWPFS
 ADLYNNKTNHPPSEDPORLUTGLVESLMFSHQPTDDCOQLQTLFTTEERKILLEA
 RNKVPADRPDLOLEIDMGPELTPRGMDYVTAERESLKYRQALVAGLRGASRP
 TNLAKVREWMQGNPEPSVLELMFAFRFTPPDPTSEAKASMAIAFICGSALDIR
 KILAELEGOELHDLVKEAKVYKKRTEEREROREREREERERERERERERER
 KILAAVEKSNRERDRPKIRSRGSGNIGNNRPLDKDCAVCKEKGHARDCPK
 KGNKGLVYALEBDK"

2728..6309

/gene="pol"

<2728..6309

/gene="pol"

/note="complete pol ORF, gag-pol is expressed as a
 polypeptide"

/codon_start=1

```

/product="polymerase"
/protein_id="CAC39617.1"
/db_xref="GI:14275839"
/translation="GRGSDPLPEPRVTLKEGQPELVDTGAKHSLLDPLQKLIKD
KSMVWGATGQGOQPMTRRVDLVGVGRVTHSFLVPECPALPLGRLLKMAQISF
EOKPEVSANNNKPTVTLTOLDDEYRLXPLVPRDQIQWILQFQPMANETAGMGIA
KQPPQVYQLKASAPVSVKQYPLSKERARGLRHHVRLIQQGLVFPVQSPMTPLP
VRKPPNDYRVDQDREVNRVODIHPTVPNPVNLCLALPQSRWTVLVLKAFECFL
RLHPTSQLPFAFWERDPCAGRTGQTLWTRLPOGKNSPTJEDBALHDLNFIQHPQ
VTLLOYVDLLLAGATKQDCLEGTKALLLESLDYGASAKMOICRETEYVYSIRG
GGORWLTAKRRRVVQVLPPTAKQVREFLGTFGLMIPGATLAALPLXPLTKRG
EESNAPHEKAFDAIKKALLSAPLALPDVTKPTLYVDERKKARVAVLTQIGKPRP
PVATLSKLDLPVAGWPICLKAIAVAALVKKDADKLTGQNTVIAVHALENTVPRP
DRWMTNARMTHYOSLLTERVTFAPRALNAPLLEPIDEPTHDOHLLIETGVK
KDLTIDLTGEMLTWFTDSSYMEGKRMAGAAVVDRTIMASSLEPGEISAKAELM
ALPOLRLAEGKSINITYDSVAFAVTAHVAGIYKQRTIMASLEPGEISAKAELM
ALHPRKALTIHGHGOKADKPIASRGOMADRAVKAQGVNLLPMTEPAREPRO
YTLDEWQPIKKIDPFSTPSTCTCTSDGKEITLHKEGLETVQOIHRLHILGTHLQOL
VRSIPYHVLRLPGVADSVNHCYPCQLVNPANPSIPEGRKRGSHPGAHMEVPTFK
PAKYGNKYLLVVDVTFSGWAEVPTKKESTVVAKKLIEIPFPGIPKVIQSDNGPA
FVAGVSOGLAKIIGIDMKLCAYRPOSSGOVERMNTIKETLTKETETGINDMALL
PEVLEFRVNTPGOGFLPYELLGGPPLAEIAFAHSADVLLSOPLESRLKALEMRO
RAWKOLREAYSGDIOYHREFOYDSYVRHAGNLEPWRKGPVLYLLTTPVAVVE
GTPWTASHVKKRPPDSCMKAEKTBPLKRLHVRVPTSVNNSS"
6185..8149
/gene="env"
6185..8149
/gene="env"
/codon_start=1
/product="envelope"
/protein_id="CAC39618.1"
/db_xref="GI:14275840"
/translation="MHPITLSRHLPIRGKPKRLKIPLSFASIAWELTSLTPQVNGK
RLVDSPNSEHKPLSLTWLTSRGLINISNGEPLSTGMEVLYCLRSVTPQLNDQA
TPPDVILAYAGPYVCPGPNNEEYCGNODPFCOMSTITSDNCKMPPVSOODRVYS
FVNPTSYNPNNGHGRKWKMOORVQKDVANKOISCSLSDLDIKISFTEKQKQENIQ
KWNAGISWGLVYTGSGRRKGSVLTIRLETMEPRVALGPKNGLAEGQPPLOEORP
SPNSDNTNNTSGVPTENITIKTGAKFLSIGAFALNSTPEAKISCSMLCLASGP
PYEGMARGKFNVTKEHRDQCTWGSQNKLLIEVSGKGCIGMVPSSHCLNHTEA
FNRTSESOVLPGDWMACNTGLPCVSLVNMOTLDFVMQIYDPRVYIEKVLV
DEVYRNRPKRPISTILAVMIGIAGVAGTAVGTCVPOQLKGLSNLIRIYTE
DLQALEKSVSNLESLSLSEVLYQNRGIDLFLKRGGLCVALLKECCCTVYDHSGAI
RDSMSKLRELERKHKKEKAGOGWFEQFNKSPWYVTLTSLALGPLVILLTLTPVC
LINFVAFVREDOVSANVILWLRQYQGLPS"
8216..8918
LTR
polyA_signal
8810..8915
BASE COUNT 2435 a 2174 c 2253 g 2056 t
ORIGIN
alignment_scores:
Quality: 2661.00 Length: 525
Ratio: 5.127 Gaps: 1
Percent Similarity: 98.857 Percent Identity: 95.619
alignment_block:
US-09-171-553b-4 x PEN293656 ..
Align seg 1/1 to: PEN293656 from: 1 to: 8918
1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrpThr 17
|||||
1153 ATGGGACAGACGGTGCAGACCCCTTACGTTGACCTCGACCATTCAGAC 1202
17 rgiValArgSerArgAlaHisAsnLeuSerValGlnValLysGly 34
|||||
1203 TGAAGTTAAATCCAGGCTCATTAATTTGTACAGTCAAGGTTAAGAAGGAC 1252
34 rotTrpGlnThrPheCysAlaSerGluTrpProThrPheAspValGlyTrp 50
|||||
1253 CTGGGACAGACTTCTGTCTCTGATGAGCCGACATCGATCTTGGATGG 1302
51 ProSerGluGlyThrPheAsnSerGluIleIleLeuAlaValLysAlaIle 67

```

```

|||||
1303 CCATCAGAGGGGACCTTTAATTCGAGATTATCCCTGGCTGTTAAGCAAT 1352
67 eIlePheGlnThrGlyProGlySerHisProAspGlnGluProTyrIleIle 84
|||||
1353 TATTTTAAAGCTGAGCCGGCTCTCATCCGATCCAGAGGCCCTATATCC 1402
84 eutThrTrpGlnAspLeuAlaGlnAspProProProThrValLysProTrp 100
|||||
1403 TTACCTGCAGAGATTGGCAGAGATCCTCCGATGGGATTAAACCTTGG 1452
101 LeuAsnLysProAlaGlyLysProGlyProArgGlnLeuAlaLeuGlyGly 117
|||||
1453 CTGATTAAGCCAAAGAACCCAGGTCGCCGAATTCTGCTCTTGAGAGAA 1502
117 sAsnLysHisSerAlaGlnLysValGluProSerSerTyrLeuProAla 134
|||||
1503 AAACAACACTCGGCTGAAAAGATCAAGCCCTC.TCCTCATATCTTACCC 1551
134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150
|||||
1552 GAGATTTCAGAGCCGCGGCTTGGCCGGAACCCCAATCTGTGCCCCACC 1601
151 ProLeuSerSerThrGlyCysCysGluGlyThr.SerAlaProProGly 167
|||||
1602 CCTTATCCGCGACAGGGCTGCGAGGGAGCCCTCGCCCTCTGAGAG 1651
167 IaProValValGluGlyProAlaAlaGlyThrArgSerArgArgGlyAla 183
|||||
1652 CTCGGGCGGTGAGAGACCTGCTGCGAGGACTCGAGAGCCGAGGGAGGCC 1701
184 ThrProGluArgThrAspGluIleAlaIleLeuProLeuAlaGlyThrTrp 200
|||||
1702 ACCCGGAGCGGACAGAGAGATCGCGCATTCGCTGGCCAGCTACGG 1751
200 YProProMetProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPhe 217
|||||
1752 CCCCTCCCAATACCGGGGGCCCAATTCAGACCCCTCAGATATGGGCCCTTT 1801
217 eSerAlaAspLeuTyrAsnTrpLysTrpAsnHisProProPheSerGlu 233
|||||
1802 CTTCGCGAGATCTCTATATTTGAAACATTAACCTATCCCTCTTCTCGAG 1851
234 AspProGlnArgLeuThrGlyLeuValGlnSerLeuMetPheSerHisGln 250
|||||
1852 GATCCCAAGCCGCTCAGCGGCTGTGGAGTCCCTTATGTTCTCTACACA 1901
250 nProThrTrpAspAspCysGlnGlnLeuLeuGlnThrLeuPheThrThr 267
|||||
1902 GCCTACTGGGATGATTGTCAACAGCTGTCAGACACTCTTCAACAACCG 1951
267 IugIuArgGluArgIleLeuLeuGlnAlaArgGlyAsnValProGlyAla 283
|||||
1952 AGGAGCGAAGAGAAATTCGTTAGAGGCTAAAGAAATGTTCCGTGGGACC 2001
284 AspGlyArgProThrGlnLeuGlnAsnGlnIleLeuMetGlyPheProLe 300
|||||
2002 GACGGCGACCCACGACGATTCGAAATGAGATTCACATGGGATTTCCCTT 2051
300 uThrArgProGlyTyrAspTyrAsnThrAlaGluGlyValArgLysSerLeu 317
|||||
2052 GACTCGCCCGCTGGGGACTATCAACACGCGTGAAGGATGAGGAGCTTGA 2101
317 ySileTyrArgGlnAlaLeuValAlaGlyLeuAlaArgGlyAlaSerArgArg 333
|||||
2102 AAATCTATCGCCAGGCTCTGCTGCGGCTCTCCGAGGCGCTTCAAGACGG 2151
334 ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGluPr 350
|||||
2152 CCCACTAATCTGGCTTAACGTAAAGAGGTGATGACAGGACGAGATGAAC 2201
350 oProSerValPheLeuGluArgLeuMetGluAlaPheArgArgPheThr 367
|||||

```

```

2202 TCCCTAGATTTTCTTGAGAGGCTCATGGAAGCCTTCAGGCGGTCACCC 2251
367 ropheaspProthrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 383
|||||
2252 CTTTGTATCTACTCGAGAGGCTCAGAAAGCCTCATGCTCGGCTTC 2301
384 ILGLYGLInSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGlu 400
|||||
2302 ATAGGACAGTCAGCCCTGATATCAGAAAGAGCTTCAGAGACTGGAAG 2351
400 YLeuGlnGluAlaGluLeuArgSpleuValArgGluAlaGluValT 417
|||||
2352 GTTACAGAGGCTGAGCTACATGATCTAGTGAAGGAGGCAAGAAAGTGT 2401
417 YTTTArgArgGluThrGluGluGluLysGluGlnArgLysGluLysGlu 433
|||||
2402 ATTACAAAGGAGGACAGAAAGAAAGGAGCAACAAAGAAAGAGAGAA 2451
434 ArgGluGluArgGluGluArgArgArgArgGlnGluLysAsnLeuTh 450
|||||
2452 AGAGAGCAAGAGGAGGAAAGACTATATAACGCAAGAGAGAAATTTGAC 2501
450 rLysIleLeuAlaIleValAlaGluGluLysSerSerArgGluArgGlu 467
|||||
2502 TAAAGATCTGGCTGCAGCTGTTGAAGGAAAGCAATAGGAAAGAGAGA 2551
467 rGAspPheArgLysIleArgSerGlyProArgGlnSerGlnLysAsnLeuGly 483
|||||
2552 GAGATTTTGAAGAAATTAGTGTCAGGCCCTAGACAGTCAGGAAACCTGGGC 2601
484 AsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysG 500
|||||
2602 AATAGGACCCCACTCGACCAAGGACCAATGTGCATATTGTAAAGAAAAAG 2651
500 YHISTPAlaArgAsnCysProLysLysGlnLysAsnLysGlyProLysVal 517
|||||
2652 ACCTGGGCAAGGAGCTGCCCAAGAGGAAACAAAGAGCTGAGAGTCT 2701
517 euAlaLeuGluGluAspLysAsp 524
|||||
2702 TACCTCTGGAAGAGATTAAGAC 2724

seq_name: gb_pat:AX052636

seq_documentation_block:
LOCUS AX052636 6076 bp DNA PAT 12-JAN-2001
DEFINITION Sequence 21 from Patent WO0071726.
ACCESSION AX052636
VERSION AX052636.1 GI:12226826
KEYWORDS
SOURCE
. porcine endogenous retrovirus.
ORGANISM porcine endogenous retrovirus
Virus; Retroviral viruses; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
REFERENCE
1 (bases 1 to 6076)
AUTHORS Federispiet,M.J.
TITLE Methods to inhibit infectious agent transmission during
xenotransplantation
JOURNAL Patent: WO 0071726-A 21 30-NOV-2000;
Mayo Medical Ventures (US)
FEATURES
Location/Qualifiers
source 1..6076
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
BASE COUNT 1613 a 1512 c 1609 g 1342 t
ORIGIN

```

```

alignment_scores:
Quality: 2659.00 Length: 525
Ratio: 5.143 Gaps: 1
Percent Similarity: 98.476 Percent Identity: 95.619

```

```

alignment_block:
US-09-171-553B-4 x AX052636 ..
Align seg 1/1 to: AX052636 from: 1 to: 6076

1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrpTh 17
|||||
1183 ATGGACAGAGGCTGACACACCCCTTATGTTGACTCTCGACCATTTGGAC 1232
17 rGluValArgSerArgAlaHisAsnLeuSerValGluValLysLysGly 34
|||||
1233 TGAAGTTAATTCAGAGGCTCATTAATTGTCTCAGTTTCAGGTTAGAGGAC 1282
34 rOTrPGLInThrPheCysAlaSerGluThrProThrPheAspValAlaTyr 50
|||||
1283 CTGGCAACACTTCTGCTCTGATAGGCCGACATTCGATTCGATTCGATTCG 1332
51 ProSerGluGlyThrPheAsnSerGluIleLeuAlaValLysAlaI 67
|||||
1333 CCATCAGAGGGGACCTTAAATTCTGACATTAATCCTGCTGTTAAAGCAAT 1382
67 eLLePheGlnThrGlyProGlySerHisProAspGlnGluProThrIle 84
|||||
1383 TATTTTTCAGACTGAGCCGCTCTCATCCCAATCAGAGAGCCATATACC 1432
84 eUTrIPGLInAspLeuAlaGluAspProProProTrpValLysProTrp 100
|||||
1433 TTACGTGCAAGATTGTGACAGAGATCTCCGACATGGGTTAAACCTTGG 1482
101 LeuAsnLysProArgLysProGlyProArgGlnLeuAlaLeuGlyLys 117
|||||
1483 CTGAATTAAGCCCAAGAAACCCAGGTCGCCGAATTCGCTCTTGGAGAGAA 1532
117 sAsnLysHisSerAlaGluLysValGluProSerSerSerThrLeuPro 134
|||||
1533 AAACAAACACTCGGCTGAAAGATCAAGCCCTC.TCCTCATATCTACCC 1581
134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150
|||||
1582 GAGATTGAGAGAGCCGCGGCTTGGCCGGAACCCCAATCTGTCCGCCAC 1631
151 ProLeuSerSerThrGlyCysCysGluGlyThr.SerAlaProGlyAla 167
|||||
1632 CCTTATCTGGACACAGGCTGCTGAGAGGAGCCCTGCTGCTGAG 1681
167 lAProValAlaGluGlyProAlaAlaGlyThrArgSerArgArgGlyAla 183
|||||
1682 CTCGGCGGTGAGAGACTGCTGACAGGACTCGAGCCGAGGCGGCC 1731
184 ThrProGluArgThrAspGluIleAlaAlaLeuProLeuArgThrTyrG 200
|||||
1732 ACCCGAGAGCGGACAGAGAGATCGCAGATTAACCGCTGCGCAGATACG 1781
200 yProPMeTrProGlyGlnLeuGlnProLeuGlnTyrTrpProPhe 217
|||||
1782 CCTCCACACCGGGGGCCCAATTCAGACCCCTCCAGATATATGGCCCTTT 1831
217 eRSerAlaAspLeuTyrAsnTrpLysThrAsnHisProProPheSerGlu 233
|||||
1832 CTTCGACGATCTATTAATGGAACCTAACCTATCCCTTCTCGGAG 1881
234 AspProGluArgLeuThrGlyLeuValAlaGlnSerLeuMetPheSerHisG 250
|||||
1882 GATCCCAACCCCTCAGCGGGGTTGTGTAGATTCCTTATGTTCTTCACCA 1931
250 nProThrTrpAspAspCysGlnGlnLeuLeuGlnThrLeuPheThrThrg 267
|||||
1932 GCTACTTGGGATGATTGTCAACAGCTGCTGACAGACACTCTTCACACCG 1981
267 lUGluArgGluArgIleLeuLeuGluAlaArgLysAsnValProGlyAla 283
|||||
1982 AGGAGCGAGAGAGATTCGTAGAGGCTAGAAAAAATGTTCCTGGGGCC 2031

```

284 AspGlyArgProThrGlnLeuGlnAsnGluLeuAspMetGlyPheProle 300
 2032 GAGGGGCGACCCAGCTGCAATGATGACATGAGATTTCCCTT 2081
 300 uThrArgProGlyTrpAspTyrAsnThrAlaGluGlyArgGlySerLeu 317
 2082 GACTGCGCCCGGTGGACTACACACGCGTCAAGAGTGGAGAGCTTGA 2131
 317 ySileTyrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgAr 333
 2132 AAATCTATCGCCAGCTCTGGTGGGGGCTCCCGGGGCGCTCAAGACGG 2181
 334 ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGluPr 350
 2182 CCCACTAATTGGCTAAGGTAAGAGAACTGATGACGAGGACCAATGAACC 2231
 350 oProSerValPheLeuGluArgLeuMetGluAlaPheArgArgPheThr 367
 2232 CCCCTCTGTTTCTTGAGAGGCTCTTGGAAGCCTTCAGGGGTACACCC 2281
 367 roPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 383
 2282 CTTTGTATCCACCTCAGAGGCGCCAAAGCCTCAGTGGCTTTGGCCCTT 2331
 384 lIleGlyInSerAlaLeuAspIleArgGlyLysLeuGlnArgGluGlu 400
 2332 ATAGGACACTGACCTTGATATTAGAAAGAAAGCTTCAGAGACTGGAAAG 2381
 400 yLeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluLysVal 417
 2382 GTTACAGAGGCTGAGTTCGTGATCTAGTGAAGGAGGACAGAAAGAT 2431
 417 yTTrpArgArgGluThrGluGluGluLysGluGlnArgGlySerGlu 433
 2432 ATTACAAAAGGAGACAGAAAGAAAGGAAAGAAAGAAAGAGAGAGA 2481
 434 ArgGluGluArgGluGluArgArgAspArgArgGlnGluLysAsnLeuTh 450
 2482 AGAGAGGAAAGGAGGAAAGAGCTAATAACCGCAAGAAAGAAATTTGAC 2531
 450 rLysIleLeuAlaIleValAlaGluGlyLysSerSerArgGluArgGlu 467
 2532 TAAGATCTTGCTGCTGAGTGTGAAGGGGAAAGCAATACGGAAGAGAGA 2581
 467 rGAspPheArgGlyLysLeuArgSerGlyProArgGlnSerGlyAsnLeuGly 483
 2582 GAGATTTTGGAAAATTAGTCAAGCCCTAGACAGTCAGGGAACTGGGC 2631
 484 AsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGlu 500
 2632 AATAGGACCCCACTGACAGAGCAACATGTGCTATTGTAAAGAAAAGG 2681
 500 yHISrTPAlaArgAsnGlyProLysLysGlyAsnLysGlyProLysVal 517
 2682 ACACGTGGCAAGGAACTGCCCAAGAGGAAACAAAGAGACTGAAGGCT 2731
 517 euAlaLeuGluGluAspLysAsp 524
 2732 TAGCTCTGGAAGAGATTAAGAC 2754
 seq_name: gb_pat:AX052647
 seq_documentation_block: 6076 bp DNA PAT 12-JAN-2001
 LOCUS AX052647
 DEFINITION Sequence 32 from Patent WO0017726.
 ACCESSION AX052647
 VERSION AX052647.1 GI:12226837
 KEYWORDS
 SOURCE porcine endogenous retrovirus.
 ORGANISM porcine endogenous retrovirus.
 Viruses; Retroviral; Retroviridae; Mammalian type C
 retroviruses; 1-Mammalian type C virus group.
 REFERENCE 1 (bases 1 to 6076)

AUTHORS Federpiet, M.J.
 TITLE Methods to inhibit infectious agent transmission during
 xenotransplantation
 JOURNAL Patent: WO 00/1726-A 32 30-NOV-2000;
 Mayo Medical Ventures (US)
 FEATURES Location/Qualifiers
 source 1..6076
 /organism="porcine endogenous retrovirus"
 /db_xref="taxon:61673"
 BASE COUNT 1619 a 1505 c 1598 g 1354 t
 ORIGIN
 alignment_scores:
 Quality: 2659.00 Length: 525
 Ratio: 5.143 Gaps: 1
 Percent Similarity: 98.476 Percent Identity: 95.619
 alignment_block:
 US-09-171-553B-4 x AX052647
 Align seg 1/1 to: AX052647 from: 1 to: 6076
 1 MetGlyInThrValIleThrProLeuSerLeuThrLeuAspHisTrp 17
 1183 ATGGAGACAGACGTGACGACCCCTTAAATTGACTCTGACCATTTGAC 1232
 17 rGluValArgSerArgAlaHisAsnLeuSerValGlnValLysGly 34
 1233 TGAAGTTAAATCCAGGGCTCATTAATTGTCAGTTCAGTTCAGTTCAG 1282
 34 roTrpGlnThrPheCysAlaSerGluTrpProThrPheAspValGlyTrp 50
 1283 CTTGGCAGACTTCTGCTGCTCTGATGCGGACATTCGATTCGATTCG 1332
 51 ProSerGluGlyThrPheAsnSerGluIleLeuAlaValLysAla 67
 1333 CCATCAGAGGGGACCTTAAATTGAGATTAATCTGCTGCTTAAAGCAAT 1382
 67 eIlePheGluThrGlyProGlySerHisProAspGlnGluProTyrIle 84
 1383 TATTTTTCAGCTGACCGGCTGCTCATATCCCAATCAGAGAGCCCTAT 1432
 84 euThrTrpGlnAspLeuAlaGluAspProProTrpValLysProTrp 100
 1433 TTACGTGCAAGATTTGGCAGAGGATCTCCGCAATGAGTAAACCTTGG 1482
 101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyLys 117
 1483 CTGAATTAAGCCCAAGAAAGCCAGGTCGCCGAATTTCTGCTCTTGA 1532
 117 sAsnLysHisSerAlaGluLysValGluProSerSerTyrLeuPro 134
 1533 AAACCAACACTCGGCTGAAAAAGTCAAGCCCTC.TCCTCATATTTACCC 1581
 134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150
 1582 GAGATTTGAGAGCGCGGCTTGGCGGAACCCCAATCTTCCGCCACCC 1631
 151 ProLeuSerSerThrGlyCysGlyGluGlyThr.SerAlaProProGly 167
 1632 CCGTATCTGGACAGGCTGCGGAGGAGCCCTGCCCCCTCTCTGAG 1681
 167 lAProValValGluGlyProAlaAlaGlyThrArgSerArgArgGlyAla 183
 1682 CTCGCGCGGTGAGAGGACCTCTGACAGGACTCGAGCGGAGGCGGCC 1731
 184 ThrProGluArgThrAspGluIleAlaIleLeuProLeuArgTrpGlu 200
 1732 ACCCGGAGCGAGACAGAGATCCGACATTTACCGTGGCCACAGTACGG 1781
 200 yProProMetProGlyGlnLeuGlnProLeuGlnTyrTrpProPhe 217
 1781 yProProMetProGlyGlnLeuGlnProLeuGlnTyrTrpProPhe 217

```

1782 CCTCCACACACGGGGGCAATTGCAGCCCTCCAGTATTGGCCCTTTT 1831
217 eSerAlaAspLeuTyrAsnTrpLysThrAsnHisProProPheSerGlu 233
1832 CTTTTCGACATCTCTATATATGAAAAACATACCATCCCTTTCGGAG 1881
234 AspProGlnArgLeuThrGlyLeuValGlySerLeuMetPheSerHisGly 250
1882 GATCCCAACAGCCCTCAGGGGTTGGTGGAGTCCCTATGTTCTCTCACCA 1931
250 nProThrTrpAspAspCysGlnGlnLeuGlnLeuThrLeuPheThrPhe 267
1932 GCTACTTGGGATGATTTGTCACACACTGCTGCAGACACTCTTCACACCG 1981
267 LngLArgGlnArgGlyLeuLeuGlnLysValArgLysAsnValProGlyAla 283
1982 AGGACGAGAGAGAATTCCTGTAGAGGCTAGAAAAAATGTTCTTGGGGCC 2031
284 AspGlyArgProThrGlnLeuGlnAsnGlyLeuAspMetGlyPheProLe 300
2032 GACGGGCGACCCAGCCAGTTCGCAAAATGACATGACATGATTTCCCTT 2081
300 uThrArgProGlyTrpAspTyrAsnThrAlaGlnLysValArgLysSerLeu 317
2082 GACTGCCCCGGTTGGGACTTACACACGCGCTGAAGGTAGGGAGACTTGA 2131
317 ySileTyrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArg 333
2132 AAATCTATGCGCAGGCTCTGGTGGGGGCTCTCCGGGGGCGCTCAAGAGG 2181
334 ProThrAsnLeuAlaLysValArgGlyValMetGlnGlyProAsnGlyPr 350
2182 CCCACTAATTTGGCTAAGTAAGAAAGTGTGCGAGGACCGAATGAACCC 2231
350 oProSerValPheLeuGlnArgLeuMetGlnAlaPheArgArgPheThrP 367
2232 CCCCTGTTTCTTGAAGAGCTCTTGAAGCCCTTCAGSGGCTACACCC 2281
367 rOPheAspProThrSerGlnLysValArgGlnLysValAlaLeuAlaPhe 383
2282 CTTTTCATCCACCTCAGAGGCCCAAAAGCTCAGTGGCTTGGCCCTT 2331
384 IleGlyGlnSerAlaLeuAspLysLysLysLeuGlnArgLeuGlnGlu 400
2332 ATAGACAGTCAAGCTTGATATTAGAAAGAGCTTCAGACACTGGAAGG 2381
400 yLeuGlnGlnLysLeuArgAspLeuValArgGlnLysValT 417
2382 GTTACAGAGGCTGACCTGATCTAGTGTAGTAGAGGCGAGAGAAAGTAT 2431
417 yTyrArgArgGluThrGlnGlnGlnLysGlnArgLysGlnLysGlu 433
2432 ATTACAAAAGGAGACAGAAAGAAAGGAAAGCAAGAAAGAAAGAGAGA 2481
434 ArgGlnGlnArgGlnLysArgArgAspArgArgGlnLysAsnLeuThr 450
2482 AGAGAGGAAAGGAGGAAAGAGCTAATAAACGCGAAGGAAAGAAATTGAC 2531
450 rLysIleLeuAlaAlaValAlaGlnLysSerSerArgLysArgGlu 467
2532 TAAATCTTGGCTGACGTGGTTGAAGGAAAGCAATACGGAAGAGAGA 2581
467 rGAspPheArgLysIleArgSerGlyProArgGlnSerGlyAsnLeuGly 483
2582 GAGATTTTAGGAAATATAGTCAAGCCCTAGACAGTCAAGGAAACCTGGGC 2631
484 AsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGlnLysGly 500
2632 AATGAGACCCCACTCGACAAAGAGCAATGTGCAATATTGTAAGAAAAAGG 2681
500 yHisTAlaAlaArgAsnGlyProLysGlyAsnLysGlyProLysVal 517
2682 ACACTGGGCAAGAGACTGCCCCAAGAAAGGAAACAAAGGACTGAAAGTCT 2731

```

```

517 euAlaLeuGlnLysAspLysasp 524
2732 TAGCTCTGGAAGAGATTAAGAC 2754

```

seq_name: gb_pat:AX052638

seq_documentation_block: LOCUS AX052638 7873 bp DNA PAT 12-JAN-2001

DEFINITION Sequence 23 from Patent WO0071726.

ACCESSION AX052638

VERSION AX052638.1 GI:12226828

KEYWORDS porcine endogenous retrovirus.

ORGANISM Viruses; Retroid viruses; Retroviridae; Mammalian type C

REFERENCE 1 (bases 1 to 7873)

AUTHORS Federspiel, M.J.

TITLE Methods to inhibit infectious agent transmission during

JOURNAL xenotransplantation

Patent: WO 0071726-A 23 30-NOV-2000;

Mayo Medical Ventures (US)

FEATURES

source 1..7873

BASE COUNT 2200 a 1914 c 1961 g 1798 t

ORIGIN

alignment_scores: Quality: 2659.00 Length: 525

Percent Similarity: 98.476 Gaps: 1

Percent Identity: 95.619

alignment_block: US-09-171-553B-4 x AX052638

Align seg 1/1 to: AX052638 from: 1 to: 7873

1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisTrpTh 17

62 ATGGGACAGACGGGTGACAGACCCCTTATTGTTGACTCTGACCATTTGAC 111

17 rGluValArgSerArgAlaHisAsnLeuSerValGlnValLysGlyPr 34

112 TGAAGTTAAATCCAGGGCTCATTAATTGTACGTTCAAGTAAGAGGAC 161

34 rOTPrGlnThrPheCysAlaSerGluTrpProThrPheAspValGlyTrp 50

162 CTTGGGAGACTTCTGTCTCTGATGATGGCCGACATTCATGTTGATGATG 211

51 ProSerGlnGlyThrPheAsnSerGluIleLeuAlaValLysAlaIle 67

212 CCATCAGAGGGGACCTTAACTTGAGATTATCTGCTGTTAAAGCAAT 261

67 eIlePheGlnThrGlyProGlySerHisProAspGlnGluProTyrIle 84

262 TATTTTCAGACTGGACCCGGCTCTCATCCCAATCAGGAGCCCTATATCC 311

84 eUTHrPrGlnAspLeuAlaGlnAspProProProThrValLysProTrp 100

312 TTACGCGGCAAGATTGGCAGAGAGATCCCGCCATGGGTTAAACCTTGG 361

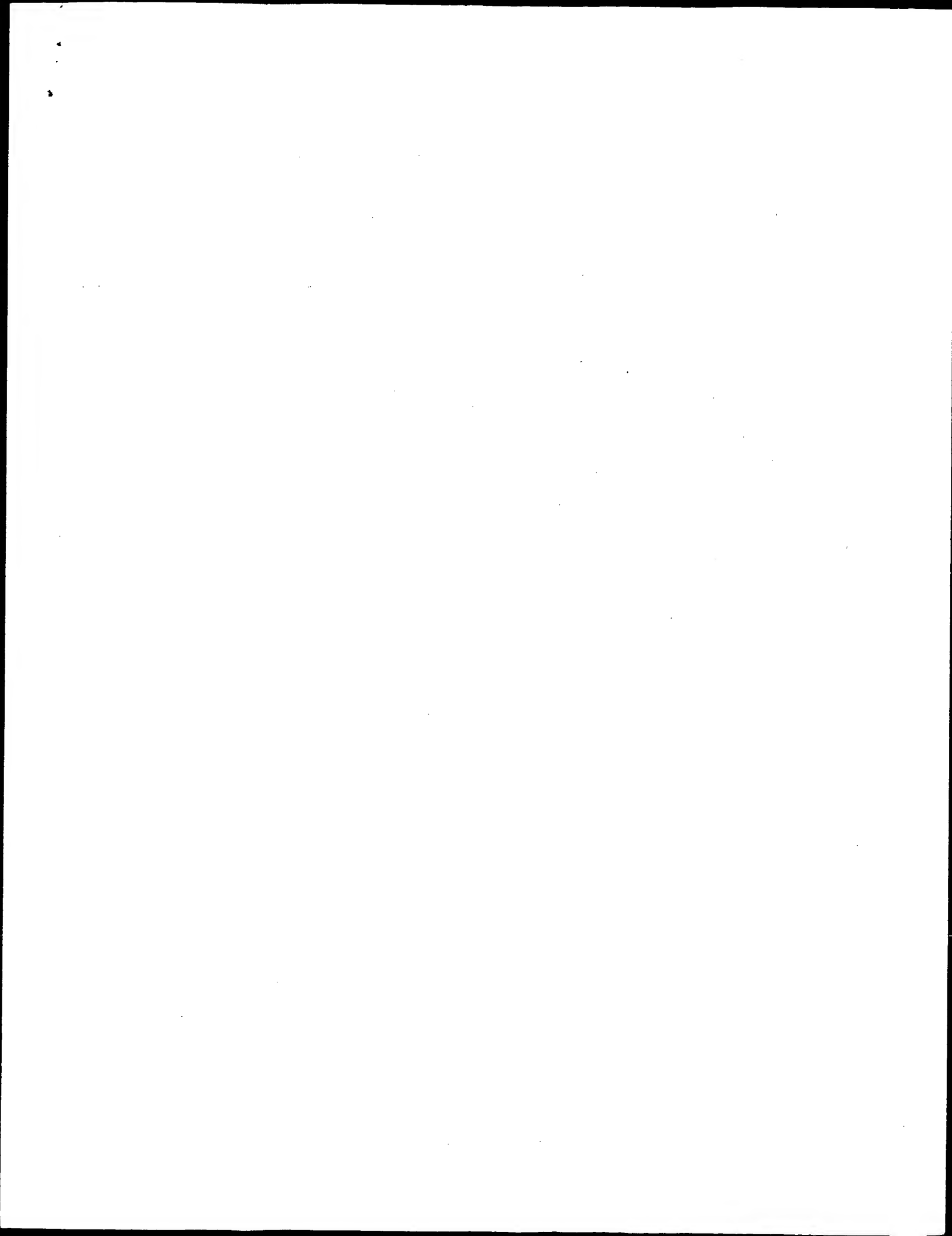
101 LeuAsnLysProArgLysProGlyProArgLysLeuAlaLeuGlyGlyLys 117

362 CTGAATTAAGCCAAAGAACGACGATCCCGCAATTTCTGGCTTGGAGAGA 411

117 sAsnLysHisSerAlaGlnLysValGluProSerSerSerTyrLeuProA 134

412 AAACAAACACTCGGCTGAAGAAAGTCAGGCCCTTC.TCCTCATATCTACCCC 460

1398 CCATCAGAGGACCTTTAATCTGAGATTCCTGGCTGTAAAGCAAT 1447
67 eilepheelnhrhglyproglyserhisproaspngluproyrilel 84
|||||
1448 TATTTTTCAGACTGACCGGCTCTCATCCCATCAGAGCGCTTATCC 1497
84 eutThrTyrGlnAspLeuAlaGluAspProProProTyrValIysProTyr 100
|||||
1498 TTACGTGCGAAGATTGGCAGAGGATCCTCCGCCATGGTTAAACCTTGG 1547
101 LeuAnIysProArgLysProGlyProArgTyrGlnLeuAlaGluGlyLys 117
|||||
1548 CTGATTAAGCCACAGAAAGCCAGGTCCTCGAATCTGGCTTGGAGAGAA 1597
117 sasnlyshisSeralagluLysValGluProSerSerSerTyrLeuProA 134
1598 AAACAACACTGCGCTGAAAAAGTCAACCCCTC.TCCCATATCTACCCC 1646
134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150
|||||
1647 GAGATTGAGGAGCCCGCGCTTGGCCGGAACCCCAATCTTCCCCCACC 1696
151 ProLeuSerSerThrGlyCysCysGlnGlyThr.SeralAProProGlyA 167
|||||
1697 CCTTATCTGGCACAGGCTGCTCCGAGGAGCCCTGCGCTCTCTGAG 1746
167 lAProValAlaGluGlyProAlaAlaGlyThrArgSerArgArgGlyAla 183
|||||
1747 CTCCGCGCGTGAGGAGACCTGCTCGAGGAGCTCGAGCCGAGGCGGCC 1796
184 ThrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrGly 200
|||||
1797 ACCCGGAGCGGACAGAGATCGCAGCATTTACCGCTGCGCAGGTACGG 1846
200 yProProMetProGlyGlyGlnLeuGlnProLeuGlnTyrTyrProPhes 217
1847 CCTCCACACACCGGGGGGCCAATTGACGCCCTCAGTATTGGCCCTTTT 1896
217 erSerAlaAspLeuTyrAsnTyrLysThrAsnHisProProPhesGlu 233
|||||
1897 CTTCTGAGATCTCTATAATTGAAAACTAACCATCCCTCTCTCGAG 1946
234 AspProGlnArgLeuThrGlyLeuValGluSerLeuMetPheSerHisGly 250
|||||
1947 GATCCCAACAGCCCTCAGCGGGTGGTGGAGTCCCTTATGTTCTCTCA 1996
250 nProThrTyrAspAspCysGlnGlnLeuGlnThrLeuPheThrThrg 267
1997 GCCTACTGGGATGATGTCACAGCTGCTGACAGACACTCTTCACAAACG 2046
267 lGluGluArgGluArgIleLeuLeuGlnAlaArgLysAsnValProGlyAla 283
|||||
2047 AGGAGCGAGAGAAATTCGTTAGAGGCTAGAAAAAATGTTCTCGGGGCC 2096
284 AspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProle 300
|||||
2097 GACGGGGGACCCACGCGAGTTGCAAAATGAGATTGACATGGGATTTCCCTT 2146
300 uThrArgProGlyTyrAspTyrAsnThrAlaGluArgLysSerLeuL 317
|||||
2147 GACTCGCGCCGCTGGGACTACAAACAGGCTGAAGGTAGGAGAGAGCTTGA 2196
317 yslIleTyrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArg 333
|||||
2197 AAATCTATCCGCAAGCTCTGGTGGCGGCTCTCGGGGCGCTCAAGACGG 2246
334 ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGluPr 350
|||||
2247 CCACATAATTGGCTAAGTAAGTAAGATGATCAGGAGCCGATATGAACC 2296
350 oProSerValPheLeuGluArgLeuMetGluAlaPheArgArgPheThrP 367
|||||
2297 CCCCCTGTTTCTTCTGAGAGGCTTTGGAAGCCTTCAGGCGGTACACCC 2346
367 ropheaSProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 383
|||||
2347 CTTTGTATCCACCTCAGAGGCCAAAGAAAGCTCAGTGGCTTTGGCTTTT 2396
384 lIleGlyInserAlaLeuAspIleArgLysLysLeuGlnArgLeuGlu 400
2397 ATAGACAGTCAGCTTGATATTAGAAAGAACCTTCAGGACTGCAAGG 2446
400 yLeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluLysValT 417
2447 GTTACAGAGAGCTGAGTTACGTGATCTAGTGAAGAGGACAGAGAAGTAT 2496
417 yTyrArgArgGluThrGluGluGluLysGlnArgLysGlyLysGlu 433
2497 ATTACAAAAGGAGACAGAGAAAGAAAGGACAAAGAAAAAGAGAGAA 2546
434 ArgGluGluArgGluLysArgAspArgArgGlnGluLysAsnLeuTh 450
2547 AGAGAGGAAAGGAGGAAAGACCTAATAAACGCAAGAGAGATTTGAC 2596
450 rLysIleLeuAlaIleValAlaGluGlyLysSerSerArgGluArgGluA 467
2597 TAAGATCTTGGCTGCAGTGGTGAAGGAAAGCAATACGAAAGAGAGA 2646
467 rGAspPheArgLysIleArgSerGlyProArgGlnSerLysAsnLeuGly 483
2647 GAAATTTTGAAGAAATTAGTCAAGCCCTTAGACAGTCAAGGAAACCTGGGC 2696
484 AsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGly 500
2697 AATAGAGACCCACTCGACAAAGGACCAATGTCATATTGTAAAGAAAAAG 2746
500 yHisTPrAlaArgAsnCysProLysLysGlyLysAsnLysGlyProLysVal 517
2747 ACAGTGGCAAGGAACTGCCCAAGAGGGAAGAAAAAGAGTGAAGGTCT 2796
517 euAlaLeuGluGluLysAsp 524
2797 TAGCTGAAAGAGATTAAGAC 2819



OM of: US-09-171-553B-5 to: GenEmbl:* out_format : pfs
 Date: Feb 24, 2002 10:10 AM
 About: Results were produced by the GenCore software, version 4.5.
 Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODE=frame+p2n.model -DEV=xlh
 -O=/cgn2.1/uspro_spool/US09171553/runat_22022002.122825.11149/app_query.fasta_1.3312
 -DB=GenEmbl -OPMT=fastap -SUFFIX=feb22std.rge -CAPOP=12.000
 -CAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blsum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTPMT=pfs
 -NORM-ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
 -USER=US09171553@cgn1_1.13817 -NCPU=6 -ICPU=3 -LONGLOG
 -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-171-553B-5
 Query length: 1194
 Database: GenEmbl:*
 Database sequences: 1472140
 Database length: 341344837
 Search time (sec): 5853.310000

score_list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
gb_pat:A66553	+ 6304.00	6651.55	0.0	8209	! A66553 Sequence 3 from Patent
gb_pat:A66552	+ 6284.00	6630.40	0.0	8196	! A66552 Sequence 2 from Patent
gb_vi:PEN133816	+ 6232.50	6575.11	0.0	8918	! AJ133816 Porcine endogenous re
gb_vi:PEN293657	+ 6229.50	6572.09	0.0	8763	! AJ293657 Porcine endogenous re
gb_vi:PEN133817	+ 6225.50	6567.77	0.0	8849	! AJ133817 Porcine endogenous re
gb_vi:PEN133818	+ 6223.50	6565.76	0.0	8750	! AJ133818 Porcine endogenous re
gb_vi:PER17013	+ 6202.50	6544.59	0.0	7808	! Y17013 Porcine endogenous reti
gb_vi:PEN293656	+ 6142.00	6479.33	0.0	8918	! AJ293656 Porcine endogenous ty
gb_pat:AR130475	+ 6099.50	6435.21	0.0	8132	! AR130475 Sequence 3 from Patent
gb_pat:AX052635	+ 6095.50	6436.69	0.0	4402	! AX052635 Sequence 2 from Paten
gb_pat:AX052634	+ 6095.50	6431.90	0.0	7362	! AX052634 Sequence 19 from Pat
gb_om:AF038600	+ 6087.50	6422.51	0.0	8132	! AF038600 Sus scrofa porcine en
gb_om:AF038599	+ 5979.50	6308.21	0.0	8132	! AF038599 Sus scrofa porcine en
gb_pat:AX052638	+ 5862.00	6184.16	0.0	7873	! AX052638 Sequence 23 from Pat
gb_pat:AX052647	+ 5846.50	5958.51	0.0	6076	! AR130473 Sequence 1 from Paten
gb_pat:AR130473	+ 5637.50	5946.35	0.0	8060	! AR130473 Sequence 32 from Pat
gb_pat:AX052636	+ 5635.50	5946.87	0.0	6076	! AX052636 Sequence 21 from Pat
gb_pat:A66551	+ 4840.00	5110.60	2.2e-276	3320	! A66551 Sequence 1 from Patent
gb_om:AF038601	+ 4899.50	4732.29	2.5e-255	7333	! AF038601 Sus scrofa porcine en
gb_pat:AR130474	+ 4489.50	4732.29	2.5e-255	7333	! AR130474 Sequence 2 from Paten
gb_htg:AC092712	+ 4433.50	4641.60	2.9e-250	214764	! AC092712 Mus musculus clone
gb_htg:AC069310	+ 4424.50	4633.83	7.8e-250	177788	! AC069310 Mus musculus chromc
gb_htg:AC068810	+ 4424.50	4633.41	8.2e-250	185872	! AC068810 Mus musculus chromc
gb_htg:AL591952	+ 4418.50	4626.03	2.1e-249	207677	! AL591952 Mus musculus chromc
gb_htg:AC079547	+ 4412.50	4620.58	4.2e-249	188530	! AC079547 Mus musculus clone
gb_htg:AC079473	+ 4412.50	4619.89	4.6e-249	203155	! AC079473 Mus musculus clone
gb_ro:AC005743	+ 4378.00	4584.22	4.5e-247	185548	! AC005743 Mus musculus chromc
gb_htg:AC093350	+ 4374.00	4579.44	8.3e-247	196755	! AC093350 Mus musculus clone
gb_om:AF151794	+ 4359.00	4592.88	1.5e-247	8431	! AF151794 Phascollartectos clereu
gb_om:AC080144	+ 4357.00	4561.19	8.6e-246	202333	! AC080144 Mus musculus chromc
gb_htg:AL592224	+ 4355.00	4558.34	1.2e-245	218938	! AL592224 Mus musculus chromc
gb_ro:MMHC438N12	+ 4344.50	4550.75	3.3e-245	149886	! AF049850 Mus musculus major
gb_pat:AR136169	+ 4339.50	4572.00	2.2e-246	8655	! AR136169 Sequence 1 from Paten
gb_vi:AF053745	+ 4339.50	4572.00	2.2e-246	8655	! AF053745 Mus dunni endogenous
gb_htg:AC073736	+ 4337.50	4540.67	1.2e-244	199883	! AC073736 Mus musculus clone
gb_ro:AC012147	+ 4331.00	4533.02	3.2e-244	217034	! AC012147 Mus musculus, clond
gb_vi:GLU06065	+ 4325.00	4556.96	1.5e-245	8379	! U60065 Gibbon leukemia virus s
gb_htg:AC015797	+ 4320.00	4521.28	1.4e-243	219294	! AC015797 Mus musculus clone
gb_vi:PCGGPE	+ 4319.00	4550.94	3.2e-245	8088	! M26927 Gibbon ape leukemia vir
gb_htg:AC074337	+ 4317.00	4523.02	1.2e-243	129317	! AC074337 Mus musculus chromc

gb_htg:AC073758
 gb_htg:AC091703
 gb_htg:AC079555
 gb_htg:AC087145
 gb_htg:AC020959

seq_name: gb_pat:A66553

seq_documentation_block:

LOCUS A66553 8209 bp DNA PAT 29-MAR-1999
 DEFINITION Sequence 3 from Patent WO9740167.
 ACCESSION A66553
 VERSION A66553.1 GI:4538106
 KEYWORDS
 SOURCE porcine endogenous retrovirus.
 ORGANISM porcine endogenous retrovirus

Viruses; Retroviral viruses; Retroviridae; Mammalian type C
 retroviruses; 1-Mammalian type C virus group.
 1 (bases 1 to 8209)
 Galbraith,D.N., Haworth,C., Lees,G.M. and Smith,K.T.

PORCINE RETROVIRUS
 PATENT: WO 9740167-A 3 30-OCT-1997;
 Q ONE BIOTECH LTD (GB)
 .Location/Qualifiers
 1. 8209

FEATURES

source
 /organism="porcine endogenous retrovirus"
 /db_xref="taxon:61673"

CDS

588..2162
 /codon_start=1
 /product="VIRAL CORE PROTEIN"
 /protein_id="CAB39347.1"
 /db_xref="GI:4538107"
 /translations="MSGVTVTPLSLTLDHWTEVRSRAHNLISVQKGPQTCASEMP
 TFDKPPSPSEGTNSEILAVKALIFQTCGSDHQDEPYILTWDLAEDPPWVKPWLN
 KPRKPGPRILALGKNNKSAEKVESSYLPRDQGAADLAGTPTCSPTLSLSCGCEG
 TSAPPAGVYVPEAATRRSGATPERTAIALPLRTYPPMPGQOLAPQWPFSS
 ADLYNWKTHPPSEDPQRLTGLVESIMFDSHQPTWDDCCQLLTFTTEERERILLEA
 RKNVPGADGPTOLQEDIMGFPLTRPGWYNTAEGRESLKIYQALVAGLGSARRP
 TNLAKYEVNMGQNEPSPVLEIRMEAFRPTPTDPTSEAKASVALATIGQSALDIR
 KKLQRLGLEQAEALDLVREAKVYRRTEEEKERKEKEREREERERDRRQKMLT
 KILAAVYEGKSSRREREDFRKIRGSRQSGNLGNRLPLDKDQCAYPEKKGHWARNCPK
 <2163..5747
 /codon_start=1
 /product="REVERSE TRANSCRIPTASE"
 /protein_id="CAB39348.1"
 /db_xref="GI:4538108"

CDS

<2163..5747
 /codon_start=1
 /product="REVERSE TRANSCRIPTASE"
 /protein_id="CAB39348.1"
 /db_xref="GI:4538108"
 /translations="GRRGSDPLPEPRVTLLKVEGQVFEVLDTGAESHVLLQPLGLKLE
 KKSWMGATGQROYPMTRTVDLGVGRVTHSPFVPECPVLLGRDLTKMAQISF
 KQPEVSVNNKPIVTLLQDDEYRLYSPQKPDQDIOSMLEQFPQANETAGMGLA
 KOVPOVIQLKASATPSVSRQYPLSREARGHPVQRLLIQOGLIVPVSQWNTPLLP
 VRKPGNDYRPVDLREVNKRQVODIHTPTPNPNLLSALUPPERNWTYVLDLKDAPFL
 RLHTQSPQFAFWRDPDGTGQTQTLTWLPGQKNSPTIFDEALRDLANFRIOHPQ
 VTLQTYVDDLLLAGATQDCLEGTKALLLSLDGLYRSASAKAQICRRREVTYLSLR
 GQSWLTERAKFTVPIPAFTAKQVREFLGTAGFCRLWPGFATLAAPLYPLTKEGK
 GFSWAEHOKAFDAIKKALLSAPALPDVTPEFTLYVDERKGAVGLTOLGPWRR
 PVAYLSKKLDPVAGWPVCLKAATAVALIVKADKLTILQNTIVIPHALENIYQPP
 DMTWTHVYOSLLTDERVTPAPALNAPFLPEETDEPVTHDCHQLLEETGVR
 KDLTDPLMTGEVLTFTDSSYVVEGKRMAGAAVDGTIRTINASSPEGTSQAQALM
 ALFQALRAEGKSTINTYTSRYAFATAHVGAIKYKQGLLTISAGRIKKEITLSLE
 ALHLPKRLAIHCPGHQKADLLSRGNOMADRVAKAQAVALNLLPIETPKAPEPRQ
 LTHQWQETIKKIDQSETPEGTCTYSGKEILPHKEGLEYVQOIHRLTHLTKHLOOL
 VRTSPYHVLRLPQADVSVKHCYQCLVANNPSRIPGRKLRGSHGAWHEVDYFEVK
 PAKYGNKYLIVVDTSQWVEATPTKRTSTVVAKKILIEIFRFGIPGIGSDNGPA
 FVAQVSGGLAKLIGIDWKLHCAYRQPSQGOVERMNETIKETIKTKITETGIDNWWALL
 PFVFLRNTGPGQITPELYLGGPPLAEAFSAHSADVLISQPLSRKLEAWVRQ
 RANKQLREAYSGDLOVPHRFQVDSYVVRHAGNLETRWKGPIYLVLTPTTAVKVE
 GIPTLWASHVKKPAPPPDSGWAERKENPLKRLHLRVPYSNNNSPGQ"
 5620..7590
 /codon_start=1
 /product="VIRAL ENVELOPE PROTEIN"
 /protein_id="CAB39349.1"
 /db_xref="GI:4538109"

CDS

/translation= "MHPTLSRRHLPTRGGEKRRRLRIPUSFASIANWPLTLITTPQASSK
RLTDSNHPRIPLSLTWLIIDPDGTGVNSTRGVAPRGVTWMPHELHFLRLINPAVKSTP
PNLVRSGYFCYCGTGEKERYKSGSGESFCRRWSCVTSDNDWKWPLISLQDRVKFSEVN
SGPGKYKMKMLYKDCSPSDDLKISFTKRTGKYSKVDKWLGLNSFLLYGGAG
SLTLIRLRETCTEPPVANGDPKVLAEOPPALEPHNLVPQLTSLRPDLITQPPSNS
TTGLITNTPRNSGPVPTGQRLFSLIQGAFOAINSTDPDATSSCWLCCLSSGPPYYE
GMAKERFNVTKHNRKNTWGRNKLTLTEVSGKTCIGKAPPSHQHLICYSTVYVEQA
SNOYLVPGYNRWMACTGLTFCVSTSVFNQKDFCVMWQIYIPRVYIHPPEVYLDYD
LYRNRKFPVSLTSLAVMLGLTAVGVTGTAALITGPOOLEKGLGELHAAAMTEDLRA
LKESVNLRESLTSLSVVLQNRGLDLFLREGGLCAALKECCCFYVDHSGAIRDMS
NKLKRLERRREREADOCWFSGWPNRSPWMTLLISALTGLPLVLVLLLLLTITVGPCLINR
FNAPVRSVAQVQIVLWVROQYOGLSQGETDL"
BASE COUNT 2168 a 2064 c 2154 g 1823 t
ORIGIN

alignment_scores:
Quality: 6304.00 Length: 1194
Ratio: 5.280 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-171-553b-5 x A66553

Align seg 1/1 to: A66553 from: 1 to: 8209

1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17
2163 GGGAGACGGGGTTCGGACCCCTCCCGAGCCGAGGTAACCTTTGAAGGT 2212
17 IGLuGlyGlnProValGluPheLeuValAspThrGlyAlaGluHiserv 34
2213 GGAGGGGAACACAGTTGAGTTCTCGTTGATACCGGAGCGAGGATTCAG 2262
34 aLeuLeuGlnProLeuGlyLysLeuLysGluLysLysSerTrpValMet 50
2263 TGCTGCTACACCAATTAGAAACATAAAGAAATAAATCTCGGTGATG 2312
51 GlyAlaThrGlyGlnArgGlnTyrProTrpThrArgArgThrValas 67
2313 GGTGCCACAGGCAACGGCAGTATCATCGACTACCGGAAGAACGGITGA 2362
67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysp 84
2363 CTTGGGAGTGGGACGGGTAAACCACTCGCTTTCTGCTCATCCCTGAGTGC 2412
84 roValProLeuLeuGlyArgAspLeuThrLysMetGlyAlaGlnIle 100
2413 CAGTACCCCTTCTAGGTAGACACTTACTACCAAGATGGGAGCTCAATT 2462
101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117
2463 TCTTTTGAACAGGAAGACCAAGAGTGTCTGTGAATAACAACCATCAC 2512
117 rValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnV 134
2513 TGTGTGACCTCCCAATTAGATGATGATATCGACTATATTCTCCCAAG 2562
134 aLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150
2563 TAAACCTTGATCAAGATATACAGTCTCGTTGGAGCAGTTTCCCAAGCC 2612
151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnVa 167
2613 TGGGCAGAAACCGCAGGATGGTTTGGCAAGCAAGTTCCCCACAGGT 2662
167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184
2663 TATTCAACTGAAGGCCAGTGTGTACACCATCATCAGTCAGACAGTACCCCT 2712
184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200
2713 TGAGTAGAGAGGCTCGAAGGAATTTGGCCGATGTTCAAGAGATTAATC 2762

201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217
2763 CAACAGGGCATCTAGTCTCTGTCCAATCCCTTGGAATACTCCCTGCT 2812
217 uproValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234
2813 ACCGGTTAGAACCCCTGGGACCAATGATTATGACCAAGTACAGGACTGA 2862
234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250
2863 GAGAGGTCAATAAAGGGTGCAGGACATACACCAACGGTCCGGAACCT 2912
251 TyrAsnLeuLeuSerAlaLeuProGluArgAsnTrpTyrThrValle 267
2913 TATAACCTCTTGAGCGCCCTCCCGCTGAACGGAAGTGTGTACACAGTATT 2962
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284
2963 GGACTTTAAAGATGCTCTTCTGCTGAGATTACACCCACTAGCCCAAC 3012
284 roLeuPheAlaPheGluTyrArgAspProGlyThrGlyArgThrGlyGln 300
3013 CACTTTTTCGCTTCAATGGAGAGATCCAGGTACGGGAAGAACCCGGGAG 3062
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317
3063 CTCACCTGGACCCGACTGCCCAAGGGTTCAAGAACTCCCGGACCATCTT 3112
317 eAspGluAlaLeuHisAspAspLeuAlaAsnPheArgIleGlnHisProG 334
3113 TGACGAAGCCCTACACAGGGACCTGGCCAACTTCAGGATCCAAACACCTC 3162
334 lnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350
3163 AGGTGACCTCTCCAGTACGTGATGACCTGCTTCTGGCGGGAGCCACC 3212
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAs 367
3213 AAACAGGACTGCTTAGAAGGTACGAAGGCACTACTGCTGGAATGTCTGA 3262
367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
3263 CCTAGGTCACAGACCTCTGCTAAGAGGCCCAAGATTCGAGGAGAGAGG 3312
384 alThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400
3313 TAACATACTTGGGTACAGTTTTCGGGGGGGGGCGAGCATGGCTGACGGAG 3362
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysGl 417
3363 GCACGGAAGAAAACTGTAGTCCAGATACCGGCCCAACCAACAGCCAAACA 3412
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
3413 AGTGAGAGAGTTTTTGGGGACACCTGGATTTTGCACACTGTGGATCCCGG 3462
434 lyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450
3463 GGTTCGGACCTTAGCAGCCCTACTACCGCTAACCAAGAAAAAGGG 3512
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467
3513 GGATTCCTCTGGGCTCTGAGCACCAGAGGCAATTTGATGCTATCAAAA 3562
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484
3563 GGCCCTGCTGAGCGACCTGCTCTGGCCCTTCCCTGACGTAACTAAACCT 3612
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500
3613 TTACCTTTTATGTGGATGACCGTAAGGGAGTAGCCCGGAGGAGTTTAACC 3662

501 GlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLe 517
|||||
3663 CAACCCCTAGGACCATGGAGAGACCTGTGCTACCTGTCAAGAAGCT 3712
|||||
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaLav 534
|||||
3713 TGATCCTGTAGCCAGTGGTTGGCCCGTAIGTCTGAAGGCTATCGCAGCTG 3762
|||||
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550
|||||
3763 TGGCCATAGTGTCAAGGACGCTGACAAATTGACCTTTGGGACAGAAATA 3812
|||||
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567
|||||
3813 ACTGTAATAGCCCCCATGCTATGGAGAACATCGTTGCGCAGCCCCCAGA 3862
|||||
567 pArgTrpMetThrAsnAlaArgMetThrHisTrpGlnSerLeuLeuLeu 584
|||||
3863 CCGATGGATGACCAACGCCCGCATGACCCACTATCAAAAGCCTGTCTCTCA 3912
|||||
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600
|||||
3913 CAGAGAGGTCACTTCGCTCCACCAGCGCTCTCAACCCCTGCCACTCTT 3962
|||||
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617
|||||
3963 CTGCCTGAAGACACTGATGAACCACTGACTCATGATTGCCATCAACTATT 4012
|||||
617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634
|||||
4013 GATTGAGGAGACTGGGGTCCGACGAGACCTTACAGACATACCGCTGACTG 4062
|||||
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650
|||||
4063 GAGAAGTCTAACCTGGTTCTACTGACGGAAGCAAGTATGCTGGTGAAGGT 4112
|||||
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667
|||||
4113 AAGAGGATGGCTGGGGCGCAGTGTGTGACGGGACCCGACGATCTGGGC 4162
|||||
667 aserSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684
|||||
4163 CAGCAGCCTGCGGAAGAACTTTCAGCGCAAAAGGCTGAGCTCATGGCCC 4212
|||||
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700
|||||
4213 TCACGCAAGCTTTGGCGCTGGCCGGAAGGAAATCCATAACATTTATACG 4262
|||||
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717
|||||
4263 GACAGCAGGTATGCCTTTGCGACTGCACACGTACACGGGGCCATCTATAA 4312
|||||
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734
|||||
4313 ACANAAGGGGGTTGCTTACTTCAGCAGGAGGGAATAAAGAACAAAGAGG 4362
|||||
734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750
|||||
4363 AARTTCTAAGCCTATTAGAGCCTTACATTTGCCAAAAGGCTAGCTATT 4412
|||||
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767
|||||
4413 ATACACTGTCTGCGACATCAGAAAGCCAAAGATCTCATATCTAGAGGGA 4462
|||||
767 nGlnMetAlaAspArgValAlaLysGlnAlaGlnAlaValAsnLeuL 784
|||||
4463 CCAGATGGCTGACCGGGTGGCAAGCAGCGACGCCAGGCTGTAAACCTTC 4512
|||||
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800
|||||
4513 TGCCCTATATAGAAACGCCCAAGCCCAAGAACCCAGACACGACGATACACC 4562
|||||
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr

4563 CTAGAAACACTGGCAAGAGATAAAAGATAGACCAGTTCTCTGAGACTCC 4612
|||||
817 oGluGlyThrCystYrThrSerTyrGlyLysGluIleLeuProHisLysG 834
|||||
4613 GGAGGGACCTGCTATACCTCATATGGAAGAAATCTCTGCCCCACAAAG 4662
|||||
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850
|||||
4663 AAGGTTTAGAATATGTCACACAGATACATCGTCTAACCCACCTAGGAAC 4712
|||||
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867
|||||
4713 AAACACCTGCAGCAGTTGGTCAGAACATCCCTTATCATGTTCTGAGGCT 4762
|||||
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884
|||||
4763 ACCAGGAGTGGTGACTCGGTGGTCAAAACATTGTGTGCCCTGCCAGCTGG 4812
|||||
884 alAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900
|||||
4813 TTAATGCTAATTCCTTCCAGAAATACCTCCAGGAAAGAGACTAAGGGGAAG 4862
|||||
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917
|||||
4863 CACCAGGGCGTCACTGGGAAGTGGACTTCACCTGAGGTAAGCGCGGTAA 4912
|||||
917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934
|||||
4913 ATACGGAACAATATCTATTGGTTTGTAGACACCTTTTCAGATGGG 4962
|||||
934 alGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950
|||||
4963 TAGAGGCTTATCTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAGAAA 5012
|||||
951 IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967
|||||
5013 ATACTGGAGGAAATTTTCCAAAGATTGGAAATACCTTAAGGTAATAGGGTC 5062
|||||
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984
|||||
5063 AGACAAATGGTCCAGCTTTCGTGCCCCAGGTAAAGTCAGGGACTGGCCNAGA 5112
|||||
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000
|||||
5113 TATTGGGSAITGATTGCAAACTGCAATGTGCATACAGACCCCAAGCTCA 5162
|||||
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017
|||||
5163 GGACAGGTAGAGGAGTGAATAAGCAATTTAAAGAGAGACCTTACCATAAT 5212
|||||
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034
|||||
5213 GACCACAGAGACTGGCATTAAATGATTGGATGGCTCTCCTGCCCTTTGTGC 5262
|||||
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050
|||||
5263 TTTTTAGGGTGAAGAACACCCCTGGACAGTTTGGGCTGACCCCTATGAA 5312
|||||
1051 LeuLeuTyrGlyGlyProProProLeuAlaGluIleAlaPheAlaHisSe 1067
|||||
5313 TTGCTACTCGGGGACCCCGCTTGGCAGAAATTTGCCATAG 5362
|||||
1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084
|||||
5363 TGCTGATGTGCTGCTTTCCAGCCTTTGTTCTTAGGCTCAAGCGCTCG 5412
|||||
1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSerGly 1100
|||||
5413 AGTGGGTGAGCAGCAGCGTGGAGCAGCTCCGGGAGGCTACTCAGGA 5462
|||||
1101 GlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTyrVa 1117
|||||

5463 GGAGACTTGCAGTTCCACATCGCTTCCAAAGTTGGAGATTGAGTCTATGT 5512
 1117 lArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProTyrL 1134
 5513 TAGAGCCACCGTGCAGGAACCTCGAGACTCGGTGGAGGACCTTATC 5562
 1134 euValLeuLeuThrProThrAlaValLysValGluGlyIleProThr 1150
 5563 TCGTACTTTTGGACCAACACGCGGTGTGAAAGTGAAGGAAATCCCCACC 5612
 1151 TrpIleHisAlaSerHisValLysProAlaProProProAspSerGlyTr 1167
 5613 TGGATCCATGATCCACGCTTAAAGCGGCGCCACCTCCGAGTCGGGTG 5662
 1167 pArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgLeu 1184
 5663 GAGAGCGGAAAGACTGAGAAATCCCTTAAGCTTCGCTCCATCGCCTGG 5712
 1184 aProTyrSerAsnAsnSerProGlyGln 1194
 5713 TTCCTTACTCTAACAAATACTCCCGAGGCCAG 5744

seq_name: gb_pat: A66552

seq_documentation_block: 8196 bp DNA PAT 29-MAR-1999
 LOCUS A66552
 DEFINITION Sequence 2 from Patent WO9740167.
 ACCESSION A66552
 VERSION A66552.1 GI:4538105

KEYWORDS porcine endogenous retrovirus.

SOURCE porcine endogenous retrovirus.

ORGANISM Viruses; Retroviridae; Retroviridae; Mammalian type C
 retroviruses; 1-Mammalian type C virus group.

REFERENCE 1 (bases 1 to 8196)

AUTHORS Galbraith, D.N., Haworth, C., Lees, G.M. and Smith, K.T.

TITLE PORCINE RETROVIRUS

JOURNAL Patent: WO 9740167-A 2 30-OCT-1997;

Q ONE BIOTECH LTD (GB)

FEATURES Location/Qualifiers

1..8196

/organism="porcine endogenous retrovirus"

/db_xref="taxon:61673"

BASE COUNT 2165 a 2061 c 2146 g 1820 t 4 others

ORIGIN

alignment_scores:

Quality: 6284.00 Length: 1194

Ratio: 5.272 Gaps: 0

Percent Similarity: 99.832 Percent Identity: 99.749

alignment_block:

US-09-171-553B-5 x A66552 ..

Align seg 1/1 to: A66552 from: 1 to: 8196

1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17

2149 GGGGAGACGGGTTCGAGCCCTCCCGAGCCGAGGTAACCTTTGAAGGT 2198

17 lGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34

2199 GGAGGGCAACACAGTTGAGTTCCTGTTGATACCGGAGCGGAGCATTCAG 2248

34 aLeuLeuGlnProLeuLysLysLeuLysGluLysLysSerTrpValMet 50

2249 TCGTGTACACCATTAGGAAACCTTAAAGAAAAAATCTCGGTGATG 2298

51 GlyAlaThrGlyGlnArgGlnTrpProThrArgArgThrValas 67

2299 GTGCGCACAGGGCAACGAGTATCCATGGACTACCCCGAAGAACCGGTGA 2348

67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysP 84
 2349 CTTGGGAGTGGGAGGGTAAACCCACTCGTTTGGTTCATCCCTGAGTGC 2398
 84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100
 2399 CAGTACCCTTCTAGGTAGAGACTTACTGACCAAGATGGGAGCTCAAAAT 2448
 101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117
 2449 TCCTTTGAACAAGGAGACAGCCAGAGTCTGTGAATAACAACCCATCAC 2498
 117 rValLeuThrLeuGlnLeuAspGluTyrArgLeuTyrSerProGlnV 134
 2499 TGTGTTGACCTCCCAATTAGATGATGAATATCGACTATATTTCTCCCAAG 2548
 134 aLysProAspGlnAspGlnSerTrpLeuGluGlnPheProGlnAla 150
 2549 TAAAGCCTGATCAAGATATACAGTCTCGTGGAGCAGTTTCCCAAGCC 2598
 151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProGlnVa 167
 2599 TGGCAGAAACCGCAGGATGGGTTTGGCAAGCAAGTTCCCCACAGGT 2648
 167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184
 2649 TATTCAACTGAAGCCAGTGTACACCACTATCAGTCAGACAGTACCCCT 2698
 184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200
 2699 TGAGTAGAGGCTCGAGAGGAATTTGGCCGCTGTTCAAGATTAATC 2748
 201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217
 2749 CAACAGGGCATCTAGTTCCTGTCCTCAATCCCTTGGAAATCTCCCTGCT 2798
 217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234
 2799 ACCGTTTAGGAAGCCTGGGACCAATGATTATCGACCACTAGTACAGACTGA 2848
 234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250
 2849 GAGAGGTCAATAAAGGGTGCAGGACATACCCCAACGGTCCCGAACCCCT 2898
 251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValle 267
 2899 TATAACCTCTTGAGCGCCTCCCGCTGACGGAAGTGGTACAGATATT 2948
 267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284
 2949 GGACTTAAAGATGCCCTTCTTCTGCTGAGATTACACCCCACTAGCAAC 2998
 284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
 2999 CACTTTTGGCTTTCGAATGGAGATCCAGGTACGGGAAGAACCGGCGAG 3048
 301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317
 3049 CTCACCTGGACCGAGTCCCAAGGGTTCAGAACTCCCGGACCATCTT 3098
 317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
 3099 TGACGAAGCCCTACACAGGACCTGGGCAACTTCAGGATCCCAACACCTC 3148
 334 lValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350
 3149 AGGTGACCTTCCAGTACGTAGTACCTGCTTCTGGCGGAGGCCACC 3198
 351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAs 367
 3199 AACAGCAGCTGCTTAGAAGGTACGAAGGCCTACTGCTGGAATTGCTGA 3248
 367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384

|||||
3249 CCTAGGCTACAGACCTCTGCTAAGAGAGGCCAGATTTCAGGAGAGAGG 3298
384 alThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTyrLeuThrGlu 400
|||||
3299 TAACATACTTGGGTACAGTTTGGGGGGGGGACCGCATGCTGACGGAG 3348
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG1 417
|||||
3349 GCACGGAAGAAACTGTAGTCCAGATACCGGCCCCCAACCACAGCCAAACA 3398
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
|||||
3399 AGTCAGAGAGTTTTTGGGGCAGCTGGATTTTGCAGACTGTGGATCCCGG 3448
434 lyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450
|||||
3449 GGTITGGACCTTACAGACGCCACTCTACCCGCTAACCAAGAAAGGG 3498
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467
|||||
3499 GGATTCCTCGGCTCTGAGCACCAGAGCATTTGATGCTATCAAAAA 3548
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysPro 484
|||||
3549 GGCCCTGCTGAGCGCACCTGCTCTGGCCCTCCCTGACGTAACATAACCT 3598
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500
|||||
3599 TTACCTTTTANGTGATGAGGTAAAGGAGTAGCCCGGAGGAGTTTTAACC 3648
501 GlnThrLeuGlyProTrpArgArgProValAlaIleTyrLeuSerLysLys 517
|||||
3649 CAAACCTTAGSACCATGAGGAGACCTGTTCCTTACCTGTCAAAGAGCT 3698
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaVal 534
|||||
3699 TGATCCTGTAGCCAGTGGTGGCCGTATGCTGAAGGCTATCGCAGCTG 3748
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550
|||||
3749 TGGCCATACGTGTCAGAGCGCTGACAAATGACTTTGGGACAGATATA 3798
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567
|||||
3799 ACTGTAATAGCCCCCATGCTATGGACAACATGCTTCGGCAGCCCCCAGA 3848
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuT 584
|||||
3849 CCGATGGATGACCAACCGCCGATGACCCACTATCAAAGGCTGCTTCTCA 3898
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600
|||||
3899 CAGAGAGGTCACTTTGCTGCCACCGCCGCTCTCAACCTCGCACCTT 3948
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617
|||||
3949 CTGCTCTGAAGAGACTGATGAACCACTGACTCATGATTGCCATCAACTTT 3998
617 uileGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634
|||||
3999 GATTGAGGAGACTGGGTCCGCAAGGACCTTACAGACATACCGCTGACTG 4048
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650
|||||
4049 GAGAGGTGCTAACCTGGTTCCTGACCGAAGACAGCTATGTGGTGAAGGT 4098
651 LysArgMetAlaGlyAlaAlaValValAspClyThrArgThrIleTrpAl 667
|||||
4099 AAGAGGATGCTGGGGCGCAGTGGTGGACGGGACCGCCGACGATCTGGGC 4148
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684
|||||

4149 CAGCAGCCTGCCGGAAGSAACTTCAGCGAAAAGGCTGAGCTCATGGGCC 4198
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700
|||||
4199 TCACGGAAGCTTTGGCGCTGGCGGAAGGAAATCCATAAACATTTATAGC 4248
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717
|||||
4249 GACAGCAGGTATGCCCTTTCGCACTGCACACGTCACGGGGCCATCTATA 4298
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734
|||||
4299 ACAAGGGGTTGCTTACCTCAGCAGGGAGGAAATAAAGAAACAAGAGG 4348
734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750
|||||
4349 AAATTCCTAAGCCTATTAGAGCCTTACATTTGCCAAAAGGCTAGCTATT 4398
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767
|||||
4399 ATACACTGCTCTGGACATCAGAAAGCAAGATCTCATATCTAGAGGAA 4448
767 nGlnMetAlaAspArgValAlaLysGlnAlaGlnAlaValAsnLeuL 784
|||||
4449 CCAGATGGCTGACCGGGTTGCCAAGCAGCAGCCAGGCTGTTAACCTTC 4498
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800
|||||
4499 TGCCCTAATAGAAAGCCCAAGCCCAAGACCCAGACAGCAGACACACC 4548
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817
|||||
4549 CTAGAAGACTGGCAAGAGATAAAAGATAGACCAGTTCTCTGAGAGTCC 4598
817 oGluGlyThrCysTyrThrSerTyrClyLysGluIleLeuProHisLysG 834
|||||
4599 GGAGGGGACCTGCTATACCTCATATGGGAAGGAAATCTGCCCCACAAG 4648
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850
|||||
4649 AAGGTTAGAAATATGTCACAGATACATCGTCTAACCCACCTAGGAAC 4698
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867
|||||
4699 AAACACCTGCAGCAGTTGGTCAGAACATCCCTTATCATGTCTGAGGCT 4748
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884
|||||
4749 ACCAGAGTGGCTGACTCGGTGGTCAACATTGTGTGCCCTGCCAGCTGG 4798
884 aIAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900
|||||
4799 TTAATGCTAATCCTTCCAGAATACCTCCAGGAAGAGACTAAGGGGAAGC 4848
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917
|||||
4849 CACCCAGGCGCTCACTGGGAAGTGGACTTCACTGAGGTAAAGCCGGCTAA 4898
917 sTyrGlyAsnLysTyrLeuValPheValAspThrPheSerGlyTrpV 934
|||||
4899 ATACGGAACAAATATCTATTGTTTTGTAGACACCTTTTCAGGATGGG 4948
934 aGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950
|||||
4949 TAGAGCTTATCCTACTAAGAAGAGACTTCAACCGTGGTGGCTAAGAAA 4998
951 IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967
|||||
4999 ATACTGGAGGAAATTTTCCAAGATTTGGAATACCTAAGGTAAATAGGGTC 5048
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984
|||||
5049 AGACAATGTGTCAGCTTTCGTTGCCCAGGTAAGTCAGGAGCTGGCCAAGA 5098

```

984 leuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000
|||||
5099 TATTGGGATTGATTGGAACCTGCTTGTGCATACAGACCCCAAGCTCA 5148
|||||
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017
|||||
5149 GGACAGGTAGAGAGATGAATAGAACCAATTAAGAGAGACCCCTACCAATT 5198
|||||
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034
|||||
5199 GACCACAGAGACTGCATTAATGATTGGATGGCTCTCCTGCCCTTTGTGC 5248
|||||
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050
|||||
5249 TTTTGGGTGAGGAACACCCCTGGACAGTTTGGCTGACCCCTATGAA 5298
|||||
1051 LeuLeuTyrGlyGlyProProLeuAlaGluIleAlaPheAlaHisSe 1067
|||||
5299 TTGCTCTACGGGGACCCCGCTGGCAGAAATTCCTTTGCACATAG 5348
|||||
1067 tAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084
|||||
5349 TGCTGATGCTGCTTTCCAGCCTTTGTTCTTAGGCTCAAGGCGCTCG 5398
|||||
1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSerGly 1100
|||||
5399 AGTGGGTGAGGAGCGGTGAAGACAGCTCCGGGAGGCGCTACTCAGGA 5448
|||||
1101 GlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTyrVa 1117
|||||
5449 GGAGACTTGGCAAGTTCACATTCGCTTCCAGTTGGAGATTCAGTCTATGT 5498
|||||
1117 lArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProTyrL 1134
|||||
5499 TAGAGCCACCGTCGAGAAACCTCGAGACTCGGTGGAGGAGGACTTATC 5548
|||||
1134 euValLeuLeuThrThrProThrAlaValLysValGluGlyIleProThr 1150
|||||
5549 TCGTACTTTTACCACACCAACCGCTGTGAAGTCTGAAAGTCTGAAGGAATCCCCACC 5598
|||||
1151 TrpIleHisAlaSerHisValLysProAlaProProAspSerGlyTyr 1167
|||||
5599 TGGATCCATGATCCAGTAAAGTGGCGGACCTCCCGACTCGGGGTG 5648
|||||
1167 pArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgLeuV 1184
|||||
5649 GAGAGCCGAAAAGACTGAGATCCCTTAAGCTTCGCCCTCCATCGCCTGG 5698
|||||
1184 alProTyrSerAsnAsnSerProGlyGln 1194
|||||
5699 TTCCTTACTCTAACAATAACTCCCGAGGCCAG 5730
|||||

```

seq_name: gb_vi:PEN133816

seq_documentation_block:

LOCUS PEN133816 8918 bp DNA VRL 07-DEC-2000
 DEFINITION Porcine endogenous retrovirus type C proviral gag, pol and env
 genes and LTR (class B, clone 33).

ACCESSION AJ133816

VERSION AJ133816.1 GI:6688945

KEYWORDS env gene; env protein; gag gene; gag protein; pol gene; pol

SOURCE

porcine endogenous retrovirus.

ORGANISM

Viruses: Retroviridae; Mammalian type C

retroviruses; 1-Mammalian type C virus group.

REFERENCE

1 (bases 1 to 8918)
 Czauderna,F., Fischer,N., Boller,K., Kurth,R. and Tonjes,R.R.
 Establishment and characterization of molecular clones of porcine
 endogenous retroviruses replicating on human cells
 J. Virol. 74 (9), 4028-4038 (2000)
 MEDLINE 20219394

REFERENCE 2 (bases 1 to 8918)

Toenjes,R.R.

Direct Submission

Submitted (04-May-1999) Toenjes R.R., Medical Biotechnology,

Paul-Ehrlich-Institut, Paul-Ehrlich-Strasse 51-59, Hessen, D-63225

Langen, GERMANY

FEATURES

source

1..8918

/organism="porcine endogenous retrovirus"

/proviral

/strain="type C"

/db_xref="taxon:61673"

/clone="33"

/note="grown in PK15 (ATCC No. CCL-33) cell line"

1..707

LTR

misc_feature

545

/note="cap site"

710..727

/note="gly4"

1154..2725

/gene="gag"

1154..2725

/gene="gag"

/function="retroviral capsid"

/product="gag protein"

join(1154..2725,2729..6313,6189..8162)

/note="viral polypeptide"

/codon_start=1

/protein_id="CAB65339.1"

/db_xref="GI:6688946"

/db_xref="GI:6688946"

/translation="MGQTVTPSLTLTDHMTVEKSRHNLSVQVKGPMQWTFCASEWP

TFDVPWSEGTNSEIILAVKALIFQSPSSHPDPEYILTQDLADDPWVKPWLIN
 TKPRGPRIILALGKNKSAKVEPSPIRYEIEBPTWPEPPVPPVPAQCAVRS
 PSAPGAPVVEGPAAGTSRRGATPERTDEIILPLRTYGPMPGGQLQLOYPFRSS
 ADLYNWKTHNPFESEDPQRLTGLVESLMSFQPTWDDQLOTLFTTERERILLER
 RKNVPGADGRTQIQNEIDMGFPLTRPGWINDTAEGRSLKTIHQALVAGLRGASRRP
 TNLAKREVWQGNPEPSPVFLERLMEAFRRPTPDPTSEAKKASVALFVIGSALDIR
 KKLQRLGQLAEALDLVREAEKYVYRRETEEEKQEKEREERERDRRQEKNT
 KILAAVVEGSSRRERDFRKIRSGPQSGNLNRTPLDKDQCAKCKEKGHWARNCPK
 KGNKPKVLTLEEDKDRGSDPLPEPRVTLKVEGQPVFLVDTSAGHSLVLOPLGL
 KEKSNWMTGATGQRYDPTVTRTVDLAVGVTHSFLVIPCVPPLGRLDILTKMAQOI
 SFEGRPSEVSNRPIVILQLDDEYRLSPQVKPOODIQSWLEQDFQWAEATAGRG
 LAKQVPQVIOLKASATPVSRQYPLCSREAREGLWPHQRLIOGILVQSPWNTPL
 LPRLPGTNDYRPVQDLREVNKRQDIHTVPNPNLSALPPERNYNTVLDLDAEF
 CLRLHPTSQPLFAFEMRDPGTGTLTQLTQLPGQFNKSPITFDEALHRLDAFRIGH
 PQTLLQVDDLLLAGATKDCLEGTKALLLELGLDGLYRASAKAQICRREVTYLGVS
 LRGGORWLTAKKTVQIIPATTAQOVREFLGTAGFCRLWIPGFATLAAPLYLTK
 KGFESWAPHQKAFDAIKKALLSAPALAPDVTKPTFLYVDERKGVARGVLTQTLGPW
 RRPVAYLSKLDLPVAGSWPCLKAIAVAAILVDADKLTGQNTITVAPHALENIVRQ
 PPRDMTNARMTHYQSLTLTERTVPFAPALNPATLLPEETDEPVDHCHOLLEETFG
 VRKDLTDPLTGEVLTWFTDSSYVVEGKRMGAADVDTHTIWSLPEGTSAQKAE
 LMALTOALRLAEGSINLYTDSRYAFATAHVGAIAIKORGLTWSAGREIKNKEILSL
 LEALHLPKRLAIHCPGHOKAKDLISRGNOMADRVAKQAQAVNLLPIETPKAPEPR
 RQYTLDEWQEIKKIDQSETPGECTYSYKEILPHKEGLEYVQOIHRLTHLGTKHLQ
 QLVRTSPYHVLRLPGVADSWCPQLVNPANPSRIPPGKRLRSGHPGAHWEVDFTF
 VKPAQNGKYLLEVEDFTSGWVEAYTKKETSTVAKKILEEIPFRFGPKVIGSDNG
 PAFYAVQSQGLAKILGIDWKILHCAIRPQSSQGVQVERMNTIKETILKTITGNDWTA
 LLPVLFVRNTPGQGLTPLYELLYGPPPLVEIASVHSDVLLSQPLFKSRLALEV
 RORAWKOLREAYSGEGDLOVPRFOVGSVVRHRHAGNLETWKGPVLLTPTAV
 KVEGISTWHTSHVKLAPPDSSGWRKENTENLKLRLHRLVPSYNNNSQCIHPTLSW
 RHLPTRGEPKRLRPLISFASIAFWLTLITPQASSKRLDSSNPHPLSLTWLIDP
 DTGTVNSTRGKVRAPRTWMPLEHFLRLINPAKSTPNLVRSYFGCCPGTKKCYC
 GDGSEFCRRWSCVTSNDGDMKWIPISQDRKFSFVNSGPKYKVMKLYDKSCSPSD
 LDYLYKISFTKQKQENTKINGMSGIVFYVYGGAGSTLRLRLETETEPVAVG
 GORLFLSQGAFQAINSTDPAATSSCWLCSSSGPPYEGMAKCKEKNYKHEHNCCTW
 QSRNKLTLTEVSGKCTGCRAPPQSHQICISVVVEQASENQILVPGYNRWACNTGL
 TPCVSTVFNQSKDFCMVQIVPRVYHPEVLDVDEYDYNRNPKEPVSULTLAVMIG
 LGTAVGVTQTAALITGPQLEKGLHAAETDLRALESVNSLSEVLSLSEVIL
 QNRGLDILLFLREGGLCAALKECCCFYVHSGAIRDSMKRLRERRRERADQGW
 FEGWFNRPMTTLLSALTGPLVVLVLLLTGPGCLINREVAFAVRSVAVQIMVLRQO
 YQGLLSQGETDL"

4179 TTACCTTTTATGTGGATGAGCGTAAGGAGTAGCCGAGGAGTGTAAACC 4228
501 GlnThrLeuGlyProTyrArgProValAlaTyrLeuSerLysLe 517
4229 CAACCCCTAGGACCATGAGGAGACCTGTTCCTACCTCTCAAGAAGCT 4278
517 uAspProValAlaSerGlyTyrProValCysLeuLysAlaLeAlav 534
4279 TCATCTGTAGCAGTGTGGCCGCTATGCTGAAGGCTATCGCAGCTG 4328
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550
4329 TGGCCATAGTGTCAAGACGCTGCAAAATTCACCTTTCGGACAGAAATA 4378
551 ThrValIleAlaProHisAlaLeuGlnAsnIleValArgGlnProProAs 567
4379 ACTGTAATAGCCCCCATGCAATGGAGAACATCGTTTCGGACAGCCCCAGA 4428
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeu 584
4429 CCGATGGATGACCAACGCCCGCATGACCCACTATCAAAAGCCTGCTCTCA 4478
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600
4479 CAGAGAGGCTCACTTTCCTCCACGAGCGCTCTCAACCCCTGCCACTCTT 4528
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617
4529 CTGGCTGAAGAGACTGATGAACCACTGACTCATGATTCGCATCACTATT 4578
617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634
4579 GATTGAGGAGAGTGGGGTCCGAGGACCTTACAGACATACCGCTGACTG 4628
634 lYgluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650
4629 GAGAAGTGTCAACCTGGTTCCTGACGGAAGCAGCTATGTGGTGAAGGT 4678
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667
4679 AAGAGGATGGCTGGGGCGGGTGTGGACGGGACCCACACGATCTGGGC 4728
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684
4729 CAGCAGCCTGCGGAAGAACTTTCAGCCGAAAGGCTGAGCTCATGGCCC 4778
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700
4779 TCACGCAAGCTTGGCGCTGGCCGGAAGGAAATCCATAAACATTTATACG 4828
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717
4829 GACAGCAGGTATGCTTTCGCACTGCACACGCTACACGGGGCCATCTATAA 4878
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734
4879 ACANAAGGGGTTGCTTACTCTCAGCAGGAGGGAATAAAGAACAAAGAGG 4928
734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750
4929 AAATTCCTAAGCCTATTAGAGCCTTACATTTGCCAAAAGGCTAGCTATT 4978
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767
4979 ATACACTGTCTGGACATCAGAAAGCCAAAGATCTCATATCTAGAGGAA 5028
767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784
5029 CCAGATGGCTGACCGGGTGGCAACGAGCGCCGAGGCTGTTAACCTTC 5078
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800
5079 TGCCTATAAGAACGCCCAAGGCCCAAGCCCAAGACCCAGACAGCTACACC 5128

801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817
5129 CTAGAGACTGGCAAGAGATAAAAGATAGACCAGTCTCTGAGACTCC 5178
817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834
5179 GGAGGGGACCTGCTATACCTCATATGGAGAGAAATCCTGCCCAAAAG 5228
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850
5229 AAGGGTTAGAAATATGTCACACAGATACGCTCACCCACCTAGGAAC 5278
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867
5279 AAACACCTGCAGCAGTGGTCACAACATCCCCATTATCATGTTCTGAGGCT 5328
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884
5329 ACCAGGAGTGGCTGACTCGGTGTCAAACATTTGTCCTGCCACTGG 5378
884 alAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900
5379 TTAATGCTAATCCTTCCAGAAATACCTCCAGGAAGAGACTAAGGGAGC 5428
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917
5429 CACCCAGGCGCTCACTGGGAAGTGGACTTCACTAGGTAAAGCGGCTAA 5478
917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTyrP 934
5479 ATACGGAACAAATATCTATTTGTTTGTAGACACCTTTTCAGGATGG 5528
934 alGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950
5529 TAGAGGCTTATCTCTACTAGAAGAGACTTCAACCGTGGTGGCTAAAAA 5578
951 IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967
5579 ATACTGGAGAAATTTTCGAGATTTGGAATACCTAAGGTAATCGGGTC 5628
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984
5629 AGCAATGGTCCAGCTTTTGTCCAGGTAAGTCAGGACTGGCCAAGA 5678
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000
5679 TATTGGGATTTGATTTGAAATCTCATTTGTCATACAGACCCCAAGCTCA 5728
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017
5729 GGACAGGTAGAGAGGATGAATAGAACCATTTAAAGAGACCCCTTACCAAT 5778
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuPhePheVal 1034
5779 GACCACAGACTGGCAATTAATGATGGATAGCTCTCTGCTGCTTGTGTC 5828
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050
5829 TTTTATAGGTTAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATGAA 5878
1051 LeuLeuTyrGlyGlyProProProLeuAlaGluIleAlaPheAlaHisSe 1067
5879 TTGCTCTACGGGGACCCCGGCTGGTGAATAATTCCTCTACATAG 5928
1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084
5929 TGCTGATGCTGCTTTCAGGCTCTGTCTCTAGGCTCAAGGCGCTCG 5978
1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSerGly 1100
5979 AGTGGGTGAGCAACGAGCGTGAAGCAGCTCCGGGAGGCGCTACTCAGGA 6028


```
1101 ...GlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValty 1116
|||||
6029 GAAGGAGACTTGCAGTTCACATCGCTTCCCAAGTGGAGATTTCAGTCTA 6078
1116 rValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProT 1133
|||||
6079 TGTTAGACGCCACCGTGCAGAAACCTCGAGACTCGGTGGGAAGGCCCTT 6128
1133 yrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIlePro 1149
|||||
6129 ATCTCGTACTTTCACACACCAACGCTGTGAAGTCGAAGAAATCTCC 6178
1150 ThrTrpIleHisAlaSerHisValLysProAlaProProProAspSerG1 1166
|||||
6179 ACCTGGATCCATACATCCAGCTTAAGCTGGCGCCACCTCCCGACTCGGG 6228
1166 yTriArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgL 1183
|||||
6229 GTGAGAGCGGAAAGAACTGAGAAATCCCTTAAGCTTCGCCTCCATCGCC 6278
1183 euValProTyrSerAsnAsnSerProGlyGln 1194
|||||
6279 TGGTCTCTACTCTAACAAATACTCCCGCCAG 6313
seq_name: gb_vi:PEN293657
seq_documentation_block:
LOCUS PEN293657 8763 bp DNA VRL 30-MAY-2001
DEFINITION Porcine endogenous type C retrovirus proviral gag gene, pol gene
and env gene, class B, clone 213.
ACCESSION AJ293657
VERSION 1
KEYWORDS env gene; envelope; gag gene; group specific antigen; pol gene;
polymerase; polyprotein.
SOURCE
ORGANISM Porcine endogenous type C retrovirus.
VIRUSES: Retroid viruses; Retroviridae; Gammaretrovirus.
1 (bases 1 to 8763)
Krach, U., Fischer, N., Czaderna, F. and Tonjes, R. R.
Comparison of replication-competent molecular clones of porcine
endogenous retrovirus class a and class b derived from pig and
human cells
J. Virol. 75 (12), 5465-5472 (2001)
21256017
REFERENCE 2 (bases 1 to 8763)
Tonjes, R. R.
Direct Submission
Submitted (31-JUL-2000) Toenjes R. R., Medical Biotechnology,
Paul-Ehrlich-Institut, Paul-Ehrlich-Str. 51-95, Langen, 63225,
GERMANY
FEATURES
Location/Qualifiers
1..8763
/organism="porcine endogenous type C retrovirus"
/proviral
/specific_host="Sus scrofa"
/db_xref="taxon:70540"
/clone="213"
/cell_lines="PK15"
/note="class B"
1..630
467
LTR misc_feature
/note="Cap site"
633..650
primer_bind
1077..2651
gene
/note="gag"
1077..2651
CDS
/note="gag"
/codon_start=1
/product="group specific antigen"
/protein_id="CAC39625.1"
/db_xref="GI:14275842"
/translators="MGOTVTTPLSLTDHWTVEKRAHNLVSQVKGKGPWOTFCSEWP
TFDVGWPEGTFSNELLAVKAIIFQTGPGSHPDQEPYILTQDLAEEPPPPWKPWLN
```

```
KPRKPGRIILALGKKNHSAKVEPSPGIYPIETEEPTWPEOPVPPPPPAQAVRG
PSAPGAPVYEGPAAGTTSRRGATPRTDELAIIPLRTYGPMPGCGQLQIWPFS
ADLYNKTNIPFSEDQRLTGLVESLMF SHQPTWDDCQLQTLTFTTEERERILL
KLVNFGADGPTQLQNEIDMGFLTRFGWDYNTAEGRESLK1YQALYAGLGRARRP
TNLAKVREVMQGNPEPFSVLERLMEAFRRFTPPDPTSEAQKASVALAF1CQALDIR
KKLQRLGLOAEALDLVRAEKVY1YRRETTEEKEQKEKEERERERRRORRQENKT
K1LAAVVEGSSRRERDRFRKIRSGPRQSGNLGNRTPLDKDQCAYPEKCKGHWARNCPK
KNGKPKVLALEEDKD"
2652..6239
/gene="pol"
<2652..6239
/feature="pol"
/note="complete pol ORF, gag-pol is expressed as a
polyprotein"
/citation=[1]
/codon_start=1
/product="polymerase"
/protein_id="CAC39626.1"
/db_xref="GI:14275843"
/translators="GRRSDPLPEPRVTLKVEGPVEELVDTCAGHSVLLQPLGKLKE
KSNWMTGATGQRPWTTRKTVDLAVGRVTHSELV1PECPVLLGRDLITKMGAIISF
EQGRPEVSVANKPITVLTQLDDEYRLYSPQVAPDQDIQSMLEQFPQAWAETAGMLA
KQVPPV1QLKASATPVSRQYPLSREAREG1WPHVQRLIQOGL1VFPVSPWNTPLLP
VRKPGTNDYRPVODLREVNRKVD1IHTVPNP1LLSALPERNWT1VLDLKDQAFCL
RLHPTSOPLFAFWRDPTGRTGOLTWTRLPQGNKSPTEFDEALHDLANF1IHPQ
VTLLQYVDDLLLAGATKQDCLESTKALLLESLGLYRASAKKAC1GCRREVTYLG1SLR
GGQWLTFAARKTVYQIPATKQVREFLGTAGFCRLW1PGFATLAAPLY1P1TKBK
EFSWAPHQKAFDAIKALLSAPALAPDVT1KPTLYVDERKVGARGV1TQTLGPPWR
PVALSKLDPVAGSNVPC1KAI1AAVALV1KDADK1LTGONITV1APHAL1EN1VQPP
DRWNTARMTHYQSL1LTERVTFAPPAALNPAT1LPETDEPVT1HCHOLL1IETGVR
KDLTD1PLTGEVLTW1TDSGVVVEGK1MAGAAV1DGT1T1WASS1P1EGTSAQK1ELM
AL1QALRLAEGK1SINI1TDSRYAFAT1AHVGA1YKQGL1TSAGRE1K1KKEEL1SLE
ALH1PKRAL1I1HCPGOK1ADL1SRGNOM1ADRAK1AAQAV1NLLP1IET1PAP1ERQ
YTLEDW1EIK1IDQ1SETPECT1CYT1SGKE1LPH1EGLEYVQ1I1HRL1TGL1KHLQ
VRTSP1HLR1PGVADSVK1HCVPC1QLVNAN1P1IP1PKR1LGR1SHPC1AHWE1DTEVK
PAKYGNK1LLV1VDFT1FSGW1EAYPTK1ETSV1VAK1LE1E1PFR1G1PKV1GSDNGPA
FVAGV1SOG1AK1ILG1DMK1LHCAV1POSS1GOVER1MNT1KET1LTK1T1EG1NDW1ALL
PVL1FRVNT1PGQ1GL1TPY1ELLY1GGP1PLVE1ASV1HSA1DVL1SQPL1SR1K1ALE1WVRQ
RANKQ1REAI1SGG1DI1QVPHR1QVGS1VYVRR1HAG1NLE1TRMK1GP1YL1VLT1TTP1AVKV
EG1STW1HAS1HVK1LAP1PDS1GWRAE1KTEN1PLK1RL1RL1H1LVP1YS1NNN1SPGQ"
6112..8085
/gene="env"
6112..8085
/feature="env"
/codon_start=1
/product="envelope"
/protein_id="CAC39627.1"
/db_xref="GI:14275844"
/translators="MHPTLSWRHLPTRGGEKPLR1PLSFAS1TAWFL1TL1TTPQASSK
RL1DSSNPHR1PLSL1TW1IIDP1RGTV1NST1RGVAP1RGT1WHP1ELH1FCL1RL1N1FAV1KSTP
PN1LVR1S1G1FCC1PTE1KE1K1YCG1SG1ES1FC1RR1M1SC1VT1S1D1G1DK1RW1P1SL1OD1R1VK1F1SFVN
SG1PKY1VMK1LY1KDK1SC1SP1D1LY1K1S1FTE1K1G1Q1EN1Q1K1W1NG1MS1GW1IV1FY1KY1GGA
GST1L1RL1ET1ET1E1PPV1AD1PK1VLA1EQ1G1F1A1N1D1P1AT1SD1P1AT1SC1W1CL1SG1G1PPY
ST1TG1L1PT1PR1NS1GP1VK1T1G1OR1L1SQ1G1F1A1N1D1P1AT1SD1P1AT1SC1W1CL1SG1G1PPY
EG1MA1K1E1GN1VT1KE1H1R1N1O1CT1W1GS1RN1K1LT1E1Y1SG1K1G1C1G1K1AP1SH1O1H1C1N1S1V1V1E1Q
ASE1NO1Y1L1V1C1YN1E1W1AC1N1GL1PC1V1ST1V1NS1K1D1C1V1A1V1Q1V1P1RY1Y1H1PE1W1V1LE1D
DY1R1N1R1K1R1E1P1SL1T1A1V1M1GL1G1T1A1V1G1T1A1L1G1T1F1Q1Q1L1E1K1G1E1G1U1H1A1M1E1DLR
ALEE1SV1N1E1ES1L1S1E1V1V1Q1N1R1G1DL1L1F1R1EG1G1C1A1L1K1E1E1C1F1Y1D1H1SG1A1R1DS
MSK1L1R1E1R1R1R1R1E1A1D1Q1G1F1E1G1W1F1N1R1P1W1MT1L1S1ALT1G1PL1V1LL1L1L1L1L1L1G1PC1LIN
RFVAF1VR1RV1SAV1Q1M1V1LR1Q1Q1GL1SQ1ETDL"
8085..8660
polyA_signal
BASE COUNT 2301 a 2200 c 2287 g 1974 t 1 others
ORIGIN
LTR
polyA_signal
8655..8660
1
alignment_scores:
Quality: 6229.50 Length: 1195
Ratio: 5.235 Gaps: 1
Percent Similarity: 99.582 Percent Identity: 98.996
alignment_block:
US-09-171-553B-5 x PEN293657
```

Align seg 1/1 to: PEN293657 from: 1 to: 8763

1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17
2652 GGGAGCGGGGTCGGAGCCCTCCCGAGCCAGGGTAACTTTGAAGGT 2701
17 lGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34
2702 GGAGGGCAACACAGTTGAGTCTCTGTTGATACCGGAGCGGAGCATTCAG 2751
34 aLeuLeuGlnProLeuGlyLysLeuLysGluLysLysSerTrpValMet 50
2752 TGCTGCTACACCATTAGGAACTTAAAGAAATAAATAATCTGGGTGATG 2801
51 GlyAlaThrGlyGlnArgGlnTyrProThrArgArgThrValAs 67
2802 GGTGCCACAGGCAACGGCAGTATCCATGGACTACCCGAAACCGTTGA 2851
67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysP 84
2852 CTTGGCAGTGGGACGGGTAACCCACTCGTTTCTGGTCACTCCCTGAGTGCC 2901
84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100
2902 CAGTACCCCTTCTAGGTAGAGACTTACTGACCAAGATGGGAGCTCAAAAT 2951
101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117
2952 TCTTTTGAACAGGAGACCAAGTCTGTGAATACAAACCCATCAC 3001
117 rValLeuThrLeuGlnLeuAspGluTyrArgLeuTyrSerProGlnV 134
3002 TGTGTTGACCCCTCAATTAGATGATGATATCGACTATATCTCCCCAA 3051
134 aLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150
3052 TAAAGCCTCATCAAGATATACAGTCTCTGTTGGAGCAGTTTCCCCAGCC 3101
151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProGlnVa 167
3102 TGGGCAGAAACCGCAGGGATGGTTGGCAAGCAAGTTCGCCCCACAGT 3151
167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184
3152 TATTCAACTGAAGCCAGTGTACACAGTATCAGTCAGACAGTACCCCT 3201
184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200
3202 TGAGTAGAGGCTCGAGAGGAATTTGGCCGATGTTCAAGATTAATC 3251
201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217
3252 CAACAGGCATCCTAGTTCCTGTCCTCAATCCCTTGGAACTACCTCCCTGCT 3301
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234
3302 ACCGGTATAGGAGCCTGGGACCAATGATTATCGACCACTACAGGACTTGA 3351
234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250
3352 GAGAGTCAATAAAGGGTCGAGGACATACACCCAAACGGTCCCGAACCCCT 3401
251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValLe 267
3402 TATAACCTCTTGAGCCCTCCCGCTGAACGGAACCTGGTACACAGTATT 3451
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284
3452 GGACTTAAAGATGCTTCTTCTGCTGAGATTACACCCCTAGCCCAAC 3501
284 roLeuPheAlaPheGluTyrArgProGlyThrGlyArgThrGlyGln 300

3502 CACTTTTGGCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCCGGCAG 3551
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317
3552 CTCACCTGGACCCGACTGCCCAAGGGTTCAAGAACTCCCGACATCTT 3601
317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
3602 TGACGAAGCCCTACACAGGACCTGGCCAACTTCAGGATCCAAACACCCTC 3651
334 lnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350
3652 AGGTGACCCCTCCAGTACGTGATGACCTGTTCTGGGGGAGCCACC 3701
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuGluLeuSerAs 367
3702 AAACAGGACTGCTTAGAAGGTACGAAGGCACCTACTGCTGGAATGTCTGA 3751
367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
3752 CCTAGGCTACAGAGCCCTCTGCTAAGAGGCCAGATTTCGAGGAGAGAGG 3801
384 alThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400
3802 TAACATACCTTGGGTACAGTTTGGGGCGGGCGAGCGATGGCTGACGGAG 3851
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG 417
3852 GCACGGAAGAAACTGTAGTCCAGATACCGGCCCCCAACACACAGCAAAAC 3901
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
3902 AGTCAGAGAGTTTGGGGACAGCTGGATTTTCAGACTGTGGATCCCGG 3951
434 lyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450
3952 GGTTCGACCTTAGCAGCCCTCTACCCGCTAACCAAGAAAGG 4001
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467
4002 GAATTCCTCTGGGCTCTGAGCAGCAGAGGCAATTTGATGCTATCAAAA 4051
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484
4052 GGCCCTGCTGAGCGCACCTGCTCTGGCCCTCCCTGACGTAACCTAACCCCT 4101
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500
4102 TTACCTTTATGTGATGAGCGTAAGGAGTAGCCCGAGGAGTTTAAACC 4151
501 GlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLe 517
4152 CAACACCTTAGGACCATGGAGAGACCTGTTGCTTACCTGTCAAGAAAGCT 4201
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaVal 534
4202 TGATCTCTGAGCCAGCTGGTGGCCCGCTGCTGCTGAAGGCTATCGCAGCTG 4251
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550
4252 TGGCCATACCTGGTCAAGGACGCTGACAAATTGACTTTGGGACAGAAATA 4301
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProAs 567
4302 ACTGTAATAGCCCCCATGTCATTGGAGAACATCGTTGGGAGCCCCCAGA 4351
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeu 584
4352 CCGATGATGACCAACGCCGCCATGACCCACTATCAAGGCTGCTTCTCA 4401
584 hrGluArgValThrPheAlaProProAlaLaLeuLeuAsnProAlaThrLeu 600
4402 CAGAGAGGGTCACTTTTCGCTCCACCGCGCTCTCAACCCCTGCGACTCTT 4451

601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617
|||||
4452 CTGCCTCAAGAGACTGATGAACCACTGACTCATGATTGCCATCAACTATT 4501
617 uileGluGluThrGlyValArgLysAspLeuThrAspIleProIeuThrG 634
|||||
4502 GATTGAGAGACTGGGGTCCGCAAGGACCTTACAGACATACCGCTGACTG 4551
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650
|||||
4552 GAGAAGTGCTAACTGCTTCCACTCACGGAAGCAGCTATGTGTGGAGGT 4601
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667
|||||
4602 AAGAAGATGCTGAGCGGGGTGGTGGACGGGACCGCACGATCTGGGC 4651
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684
|||||
4652 CACAGCCTGCCGGAAGGAACCTTCAGCGCAAAAGGCTGAGCTCATGGCCC 4701
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700
|||||
4702 TCAGCGAAGCTTTGGCGGTGGCCGAAGGGAATCCATAACATTTATACG 4751
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrIly 717
|||||
4752 GACAGCAGGTATGCTTTGGCACTGCACACAGTACACGGGGCCATCTATAA 4801
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734
|||||
4802 ACAAGGGGGTTCCTTACCTTCAGCAGGGGGAATAAAGAACAAAGAGG 4851
734 luileLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750
|||||
4852 AAATTTCTAAGCCTATTAGAAGCCTTACATTTGCCAAAAGGCTAGCTATT 4901
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767
|||||
4902 ATACACTGCTCGACATCAGAAAGCCAAAGATCTCATATCTAGAGGGAA 4951
767 nGlnMetAlaAspArgValAlaLysGlnAlaGlnAlaValAsnLeuL 784
|||||
4952 CCAGATGGCTGACCGGGTTGCCAAGCAGCGAGCCAGGCTGTTAACCTTC 5001
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnItyrThr 800
|||||
5002 TGCTTATAATAGAAACGCCCAAGCCCAAGACCCAGACAGCAGTACACC 5051
801 LeuGluAspTrpGlnIleLysLysIleAspGlnPheSerGluThrPr 817
|||||
5052 CTAGAAGACTGGCAGAGATATAAAGATAGACAGTCTCTGTGAGACTCC 5101
817 oGluGlyThrCysTyrThrSerTyrClyLysGluIleLeuProHisLysG 834
|||||
5102 GGAGGGACCTGCTATACCTCATATGGGAAGGAATCTGCCCCCAAAAG 5151
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850
|||||
5152 AAGGGTTAGAAATATGTCACACAGATACATCGTCTCACCCACCTAGGAAT 5201
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867
|||||
5202 AAACACCTGCAGCAGTTGGTTCAGAACATCCCTTATCATGTTCTGAGGCT 5251
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884
|||||
5252 ACAGGAGTGGCTGACTGGTGGTCAACATTTGTTGGCCCTGCCAGCTGG 5301
884 aLsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900
|||||
5302 TTAATGCTTAATCTTCCAGAAATACCTCCAGGAAGAGAGACTAAGGGGAAGC 5351

901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLys 917
|||||
5352 CACCCAGGCGCTCACTGGGAAGTGGAGCTTCACTGAGGTAAAGCCGGCTAA 5401
917 STyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934
|||||
5402 ATACGGAAACAATATCTATTGGTTTTGTAGACACCTTTTCAGGATGGG 5451
934 aGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950
|||||
5452 TAGAGCTTATCTACTAAGAAGAGACTTCAACCGTGGTGGCTAAAAAA 5501
951 IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlyse 967
|||||
5502 ATACTGGAGGAAATTTTCCGAGATTGGGAATACCTAAGTAATCGGGTC 5551
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984
|||||
5552 AGACAATGGTCCAGCTTTGTTGCCCGAGGTAAAGTCAAGGACTGGCCAAGA 5601
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000
|||||
5602 TATTGGGATTCATTGGAAACTGCATTGTGCATACAGACCCCAAGCTCA 5651
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017
|||||
5652 GCACAGGTAGAGAGGATGAATAGAACCATTAAGAGAGACCTTACCAAAAT 5701
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034
|||||
5702 GACCACAGAGACTGGCATTAAATGATTGGATAGCTCTCCTGCCCTTTGTGC 5751
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050
|||||
5752 TTTTGTAGGGTTAGGAACACCCCTGGACAGTTTGGCTGACCCCTATGAA 5801
1051 LeuLeuTyrGlyGlyProProProLeuAlaGluIleAlaPheAlaHisSe 1067
|||||
5802 TTGCTCTACGGGGGACCCCGCTTGGTGAATAATTTGCTTCTGTACATAG 5851
1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084
|||||
5852 TGCTGATGTGCTGCTTTCGCCAGCCTTTGTTCTTAGGCTCAAGGCGCTCG 5901
1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSerGly 1100
|||||
5902 AGTGGGTGAGCAAGCAGCGTGGAGAGCAGCTCCGGGAGGCTACTCAGGA 5951
1101 ... GlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTy 1116
|||||
5952 GAAGGAGACTTGCAGATTCCACATCGCTTCCAAGTGGGAGATTCAGTCTA 6001
1116 rValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProT 1133
|||||
6002 TGTTAGACGCCACCATGCAGGAAACCTCGAGACTCGGTGGGAAGGCGCCTT 6051
1133 yrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIlePro 1149
|||||
6052 ATCTCGTACTTTTGACCCACACCAAGCGCTGTGAAGTGAAGTGAAGGAATCTCC 6101
1150 ThrTrpIleHisAlaSerHisValLysProAlaProProProAspSerGl 1166
|||||
6102 ACCTGGATCCATGCATCCACGTTTAAAGTGGCGCCACCTCCCGACTCGG 6151
1166 yTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgL 1183
|||||
6152 GTGGAGAGCCGAAAGAGACTGAGAAATCCCTTTAAGCTTTCGCTTCATGCC 6201
1183 euValProTyrSerAsnAsnSerProGlyGln 1194
|||||
6202 TGGTTCCTTACTCTAACAAATTAACCTCCCGAGGCCAG 6236

```

seq_documentation_block:
LOCUS   PEN133817      8849 bp      DNA           VRL           07-JAN-2000
DEFINITION Porcine endogenous retrovirus type C proviral gag, pol and env
genes and LTR (Class A, clone 42).
ACCESSION AJ133817
VERSION   1 (bases 1 to 8849)
KEYWORDS  env gene; env protein; gag gene; gag protein; pol gene; pol
protein;
SOURCE   porcine endogenous retrovirus.
ORGANISM Porcine endogenous retrovirus.
REFERENCE 1 (bases 1 to 8849)
AUTHORS  Czauderna,F., Fischer,N., Boller,K., Krach,U., Kurth,R. and
Toenjes,R.R.
TITLE    Molecular Characterization of Human-tropic and
Repllication-competent Porcine Endogenous Retroviruses
JOURNAL  Unpublished
REFERENCE 2 (bases 1 to 8849)
AUTHORS  Toenjes,R.R.
TITLE    Direct Submission
JOURNAL  Submitted (04-MAY-1999) Toenjes R.R., Medical Biotechnology,
Paul-Ehrlich-Institut, Paul-Ehrlich-Strasse 51-59, Hessen, D-63225
Langen, GERMANY
FEATURES
source
1..8849
Location/Qualifiers
/organism="porcine endogenous retrovirus"
/proviral
/strain="type C"
/db_xref="taxon:61673"
/clone="42"
/note="grown in PK15 (ATCC No. CCL-33) cell line"
1..668
506
misc_feature
/note="cap site"
671..688
/note="gly4"
1115..2686
/note="gag"
1115..2686
/note="gag"
1115..2686
/note="gag"
/function="retroviral capsid"
/product="gag protein"
join(1115..2686,2690..6271,6150..8132)
/note="viral polypeptide"
/codon_start=1
/protein_id="CAB65340.1"
/db_xref="GI:6688948"
/translat="MGOTVVTPLSLTDHWTVEVRSRAHNLVSVKKGPWOTFCASEWP
TFDVGSEGFNSSEIILAVKAIIFQTGSGHPDQEPYILTWDLAEDPPWVKPWLN
KPRKPGRIILALGEKNKHSAEKVEPSRIYETEEPTWPEQPPVPPYPAQAVRG
PSAPGAPVVEGPAAGTRSRGATPERTDEIATLPLRTYGPMPGQLOPLOWPFS
ADLYNKNTHNPPESEDPQRLTGLVESLMSHQPWDDCOQLLOTLETFEREILLLEA
RKNVPGADGRTQLQNEIDMGFLTRPGWDYNTAEGRESKIIYRQALVAGLGRARRP
TNLAKREVMOGNEPPEVSEFLRMEAFRRFTPDPTSEAQASVALATIGQSALDIR
KILQRLGLEAEALDLVREAEKVIYRTEEEKRQREEREERDRQKRLT
KILAAVVEGSSRRERDRFRKIRSGPRQSGNLGNRTPLQKDCAYCKEKGHWARNCPK
KGNKGKPVLALEDKDCGRGSDPLPEPRVTLKVEGQPVRELVDTGAEHSLVLLQPLGL
KEKSKSWATGQRPYTRTRVTLGVEGVTSHFLVPECPVPLGLRDLTKMGQAI
SFGQREPEVSVNNKPIVTLTLDDEIRYISPOVKEDQDQIQSLMEQFPQAAWETAGM
LAKQVPQVILQKASATPVSVRYPLISAREGIMPHVORLIQGLILVPPVQSPWNPPL
LPVRKPTNDYRVQDLREYNKRVQDIHTVNPYNLLSALPERNWTYVLDLKDAFF
CLRIHPTSPDLFAFEWRDPTGTGTGTWTRLPQGEKNSPTIFDEALHRLDLNFKIQH
POVTLLOYVDLLLAGATKDCLEGTAKILLESLDLGYSRASAKKAICIRREVTLYGS
LRGQPHLTERAKTVVQIPAPTAQVREFLGTAGECLWIPGFATLAAPLYPLTKE
KGFENWAPHQKAFDAIKKALLSAPALAPDVTKPTLYVDERKGVARGVTLQTLGPV
RRPVYLSLKKLDVPSWGVCLKAAVALLVKDKDLTLGQNTIVIAHALENIVRQ
PPDRMNTNMTYHQSLLTERVTFAPPALNATLLPEETDEPETHDCHQLLIETG
VRKDLTPIPTGLVLTFTDCSSVYVVEGRMAGAAVVDGTRTTIASSLPEGSAKAE
LMALTOALRAEGKSYNIYDSRYAFATAHVGAITYKQGLLTSAKREIKNKEEILSL
LEALHLPKRLAITHCPCHOKAKDLISRGNMADRVAQAAQVNLPIETTPKAPPR
ROYTLEDMQEIKKIDQFSETPGTCYTSYKELLPHKLEGLEYVQOIHRLTHLTKHLQ

```

```

QLVRTSPYHVLRLPGVADSVVVKHCVPQCLVNPANSPRIIPGKRLRGSHGHAWEVDFTIE
VKPAKYGNKYLLVFDVTFSGWEAYPTKKTSTVVAKKILDEEIPFRGIPKIVGSDNG
PAFVAQVSGQAKILGIDMKLHCAYRQSQGVEMRTIKETLTKLTETGINDMMA
LLPVLFRVNTPGFGLTYVELLYGGPPLVEIASVHADVLLSQPLFSRLKALEWV
RQRAWKQLREAYSGEGDQVPHRFQVGDVSVVRHAGNLETRWKGPVLVLLTPTTAV
KVERIPWIHSHVKKPAPPPDSGKAEKTEPNKLRLHRVVPYVSNSSSMHPTLSRR
HLPIRGKPKRLKIPLSFASIAWELTSLITPOVNGKRLVDSPNHKKLSLTLWLLTDSG
TCININSTQGEAPLGTWMPELYVCLRSVIFGLNQDQATPPDLRAIFYVCPGPNNEE
YCGNPDQFCKQWSCITNDGNMKWPVSQQDRVSYSFNNTPSYQNFYHGHRWKDQO
QVOKDVNRKQISCHSLDLDYKISFETKQENIKQWINGISWGIYVYGGSGRKGGS
VILIRLETOMEPPEVAGPNKLAEGOPPIQEOQPPSPNPSDYNTTSGSVTEPNITI
KITAKLFSLIQGAQALNSTTPEATSSCHLCLASGPPYEGMARGKENVTKERDQC
TWGSONKLTITFVSGKGCICWPPSHQHLNHTFAFNKRSQYLLVPGDVRWWACNT
GLTGVACVGTGTAAJITGPQOLEKSLNLRIVTEDLOALEKSVSNLESLSLSEAV
VLQNRGLDLFLKLEGICVAKKECCFYVDHSGAIRDMSKLRERLERREREADQ
GWFEQFNRSPPMTTLLSALTGPLVLLLLLTGVGCLINREVAFRVRSVAQIMVLR
QOYQGLLSQGETDL"
2690..6271
/mat_peptide
/gene="pol"
/function="retroviral protease and polymerase"
/product="Pol protein"
2690..6271
/gene="pol"
/mat_peptide
6150..8129
/gene="env"
/function="retroviral envelope"
/product="Env protein"
6150..8129
/gene="env"
LTR
8182..8849
polyA_signal
8742..8747
BASE COUNT 2355 a 2191 c 2281 g 2022 t
ORIGIN

alignment_scores:
Quality: 6225.50      Length: 1192
Ratio: 5.245          Gaps: 1
Percent Similarity: 99.581 Percent Identity: 99.161

alignment_block:
US-09-171-553b-5 x PEN133817

Align seg 1/1 to: PEN133817 from: 1 to: 8849

1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17
2690 GGGAGACGGGGTTCGGACCCCTCCCGAGCCAGGGTAACCTTTGAAGGT 2739

17 lGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34
2740 GGAGGGCAACCAAGTTCAGTTCCTGTTGATACCGGAGCGGAGCATTCAG 2789

34 allLeuLeuGlnProLeuGlyLysLeuLysGluLysSerTrpValMet 50
2790 TGTGCTACACCATTAGGAAATAAAGAAAAAATCCTGGGTGATG 2839

51 GlyAlaThrGlyGlnArgGlnTyrProTrpThrArgArgThrValAs 67
2840 GTGCCACAGGGCAACGCGAGTATCCATGACATCCCGAAGAACCGTTGA 2889

67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysP 84
2890 CTTGGGAGTGGGAGGGGTAAACCCACTCGTTCTTGGTCATCCCTGAGTGC 2939

84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100
2940 CAGTACCCCTTCTAGGTAGAGACTTACTGACCAAGATGGGAGCTCAATT 2989

101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117
2990 TCTTTTGAACAAGGAGACCAAGAGTCTCTGTGAATAACAACCAACATCAC 3039

```

117 rValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnV 134
|||||
3040 TGTGTGACCTCCAAATAGATGATGAATATCGACTATATCTCCCCAAG 3089
134 allYsProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150
|||||
3090 TAAAGCCTGATCAAGATATACAGTCCCTGGTTGGAGCAGATTTCGCCAAGCC 3139
151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnVa 167
|||||
3140 TGGCGAAGAACCGCAGGATGGGTTTGGCAAGCAAGTTCCTCCACAGGT 3189
167 lileGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184
|||||
3190 TATTCAACTGAAGGCCAGTCTACACAGATATCAGTCAGACAGTACCCCT 3239
184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200
|||||
3240 TGAGTAGAGAGGCTCGAGAGGAATTTGGCCGCATCTTCAAGATTAAATC 3289
201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217
|||||
3290 CAACAGGGCATCTAGTCTCTGCCAATCCCTTGGAACTACTCCCTTGCT 3339
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234
|||||
3340 ACCGGTTAGGAAGCCTGGGACCAATGATTATCGACCAGTACAGACTTGA 3389
234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250
|||||
3390 GAGAGGTCAATAAAGGGTGCAGGACATACACCAACGGTCCCGAACCCCT 3439
251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValLe 267
|||||
3440 TATAACCTCTTGAGCGCCCTCCCGCTGAAACGGAACTGGTACACAGTAT 3489
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284
|||||
3490 GGACTTAAAGATGCCCTTCTCTGCCTGAGATTACACCCCACTAGCCAAC 3539
284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
|||||
3540 CACTTTTGCCCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCGGGCAG 3589
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317
|||||
3590 CTCACCTGGACCCGACTGCCCCAAGGGTTCAAGAACTCCCCGACCATCTT 3639
317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
|||||
3640 TGACGAAGCCCTACACAGGACCTGGCCCAACTTCAAGATCCAAACACCCTC 3689
334 lnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350
|||||
3690 AGGTGACCTCTCCAGTAGTGGATGACCTGCTTCTGGCGGAGCCAC 3739
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuGluLeuSerAs 367
|||||
3740 AAACAGACTGCTTACAGGTACGAAGGCACCTACTGCTGGAATTGTCTGA 3789
367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
|||||
3790 CCTAGGCTACAGGCTCTGCTAAGAAGGCCAGATTTGCAGGAGAGAGG 3839
384 alThrTyrLeuGlyTyrSerLeuArgGlyGlnArgTrpLeuThrGlu 400
|||||
3840 TAACATACTTGGGGTACAGTTTGGGGGGGGGCGAGCATGGCTGACGGAG 3889
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysGl 417
|||||
3890 GCACGGAAGAAACTGTAGTCCAGATACCGGGCCCCAACACAGCCCAACA 3939

417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
|||||
3940 AGTGACAGAGATTTTGGGACAGCTGGATTTTCGACACTGTGTGATCCCCG 3989
434 lyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450
|||||
3990 GGTTCGACCTTAGCAGCCCCACTCTACCCGCTAACCAAGAAAGG 4039
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467
|||||
4040 GAATTCCTCTGGGCTCTCGAGCACCAGAAAGGCAATTGATGCTATCAAAA 4089
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484
|||||
4090 GGCCCTGCTGAGCGCACCTGCTCTGGCCCTCCCTGACGTAACCTAAACCT 4139
484 heThrLeuTyrValAspGluArgLysGlyValAlaAlaArgGlyValLeuThr 500
|||||
4140 TTACCCCTTTATGTTGATGAGCGTAAGGGAGTAGCCCGAGGAGTTTAAAC 4189
501 GlnThrLeuGlyProTrpArgProValAlaIleTyrLeuSerLysLysLe 517
|||||
4190 CAACCCCTAGGACCATGGAGGAGACTGTGCTTACCTGTCAAGAAAGCT 4239
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaLav 534
|||||
4240 TCATCCTGTAGCCAGTGGTTGCCCGTATGCTGGAAGGCTATCGCAGCTG 4289
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550
|||||
4290 TGGCCATACTGTCAGAGGACGCTGACAAATTGACTTTGGGACAGAATATA 4339
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567
|||||
4340 ACTGTAATAGCCCCCATGATTGGAGAACATCGTTGGGACGCCCCAGA 4389
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeu 584
|||||
4390 CCGATGGATGACCAACGCCCGCATGACCCACTATCAAGACCTGCTTCTCA 4439
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600
|||||
4440 CAGAGAGGTCACTTTGCTCCACCAGCGCTCTCAACCCCTGCCACTCTT 4489
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617
|||||
4490 CTGSCCTGAAGAGACTGATGAACAGTACTCATGATTGCCATCAACTATT 4539
617 uileGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634
|||||
4540 GATTGAGGAGACTGGGGTCCGCAAGGACCTTACAGACATACCGCTGACTG 4589
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650
|||||
4590 GAGAAGTGTAACTTGGTTCACTACGGAAGCAGCTATGTGTGGAAGGT 4639
651 LysArgMetAlaGlyAlaValValAspGlyThrArgThrIleTrpAl 667
|||||
4640 AAGAGGATGGCTGGGGGGCGAGTGGTGGACGGGCCCGCAGCATCTGGGC 4689
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684
|||||
4690 CAGCAGCCTGCCGGAAGGAACTTCAGCCCAAAAGGCTGAGCTCATGGGCC 4739
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700
|||||
4740 TCACCAAGCTTTGGGCTGGCGGAAGGAAATCCATAACATTTATACG 4789
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717
|||||
4790 GACACAGGATGCTGCTTTGCGACTGCACACGTACACGGGGGCATCTATAA 4839
717 sglnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734

|||||
4840 ACAAGGGGGTTCCTTACCTCAGCAGGGGAGGAAATAAGAACAAAGAGG 4889
734 luileuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750
4890 AAATTCCTAAGCCTATTAGAGCCTTACATTTGCCAAAAGGCTAGCTATT 4939
751 ileHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767
4940 ATACACTGTCTGGACATCAGAAAGCCAAAGATCTCATATCTAGAGGAA 4989
767 nGlnMetAlaAspArgValAlaLysGlnAlaGlnAlaValAlaLeu 784
4990 CCAGATGGCTGACCGGTTGCCAAGCAGGACGCCAGGCTGTAACTTC 5039
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnThr 800
5040 TGCTATAATAGAAAGCCCAAGCCCAAGACCCAGACGAGTACACC 5089
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817
5090 CTAGAGACTGGCAGAGATATAAAGATAGACAGTCTCTGAGACTCC 5139
817 oGluGlyThrCysTyrThrSerTyrGlyLysGlnIleLeuProHisLysG 834
5140 GGAGGGGACCTGCTATACCTCATATGGAAGGAAATCCTGCCCCACAAAG 5189
834 luGlyLeuGluThrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850
5190 AAGGGTTAGAAATGTCCAACAGATACATCGTCTCACCACCTAGAACT 5239
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867
5240 AAACCTCGACAGCTGGTGTGAGACATCCCTTATCATGTTCTGAGGCT 5289
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884
5290 ACCAGAGTGGCTGACTCGGTGTCAACATTTGTGCTCCCTGCCAGCTGG 5339
884 aAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900
5340 TTAATGCTAATCTTCCAGAAATACCTCCAGAAAGAGACTAAGGGGAAGC 5389
901 HisProGlyAlaHisTrpGluValAlaAspPheThrGluValLysProAlaLy 917
5390 CACCCAGGCGCTCAGTGGGAAGTGGACTTCACTGAGGTAAAGCGGCTAA 5439
917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934
5440 ATACGGAAACAATAATCTATTGTTTGTGACACACCTTTTCAGGATGG 5489
934 aGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950
5490 TAGAGCTTATCTACTAGAAAGAGACTTCAACCGTGTGCTGCTAAAGAA 5539
951 IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967
5540 ATACTGGAGGAAATTTTCCGAGATTGGAATACCTAAGTAATCGGGTC 5589
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984
5590 AGCAATGGTCCAGCTTCTGTCGCCAGGTAAAGTACGGACGTGGCCAGA 5639
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000
5640 TATTGGGATTGATGGAAATGCAATGTGATACAGACCCCAAGCTCA 5689
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017
5690 GGACAGGTAGAGAGTGAATAGAACCTTAAAGAGACCTTACCATAATT 5739
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034
|||||

5740 GACCACAGAGACTGGCATTAATGATTGGATGGCTCTCCTCCCTTTGTGC 5789
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050
5790 TTTTATAGGTGAGGAAACACCCCTGGACAGTTTGGGCTGACCCCTATGAA 5839
1051 LeuLeuTyrGlyGlyProProLeuAlaGluIleAlaPheAlaHisSe 1067
5840 TTGCTCTACGGGGGACCCCGCTTGGTAGAAATTCCTCTACATAG 5889
1067 rAlaAspValIleLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084
5890 TGCTGATGTGCTCTTCCAGCCTTTGTCTCTAGCTCAAGCGCTCG 5939
1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSerGly 1100
5940 AGTGGGTGAGGCAACGAGCGTGGAAAGCAGCTCCGGGAGGCTTACTCAGGA 5989
1101 ...GlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTy 1116
5990 GAAGGAGACTTGCAAGTTCCACATCGCTTCCAAGTTGGAGATTCACTTA 6039
1116 rValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProt 1133
6040 TGTTAGACGCCACCGTGCAGAAACCTCGAGACTCGGTGGAAGGCCCTT 6089
1133 yrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIlePro 1149
6090 ATCTCGTACTTTTGACCACACCAACGCGTGTGAAAGTCGAAAGAATCCCC 6139
1150 ThrTrpIleHisAlaSerHisValLysProAlaProProProAspSerGI 1166
6140 ACCTGGATTCATGCATCCCGTAAAGCGCGCCACCTCCCGATTCCGG 6189
1166 yTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgL 1183
6190 GTGGAAGCCGAAAGACTGAAATCCCTTAAGCTTCGCTCCATCCGCG 6239
1183 euValProTyrSerAsnAsnSer 1191
6240 TGGTTCTTACTGTCAATAACTCC 6265

seq_name: gb_vi:PENI33818

seq_documentation_block:
LOCUS PENI33818 8750 bp DNA VRL 14-APR-2000
DEFINITION Porcine endogenous retrovirus type C proviral gag, pol and env
genes and LTR (class B, clone 43).
ACCESSION AJ133818
VERSION AJ133818.1 GI:6688949
KEYWORDS env gene; env protein; gag gene; gag protein; pol gene; pol
protein.
SOURCE porcine endogenous retrovirus.
ORGANISM porcine endogenous retrovirus
Virus; Retrovirdae; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
REFERENCE 1 (bases 1 to 8750)
AUTHORS Czauderna, F., Fischer, N., Boller, K., Kurth, R. and Tonjes, R.R.
TITLE Establishment and characterization of molecular clones of porcine
endogenous retroviruses replicating on human cells
J. Virol. 74 (9), 4028-4038 (2000)
MEDLINE 20219394
REFERENCE 2 (bases 1 to 8750)
AUTHORS Tonjes, R.R.
TITLE Direct Submission
JOURNAL Submitted (04-MAY-1999) Tonjes R.R., Medical Biotechnology,
Paul-Ehrlich-Institut, Paul-Ehrlich-Strasse 51-59, Hessen, D-63225
Langen, GERMANY
FEATURES
source Location/Qualifiers
1. .8750
/organism="porcine endogenous retrovirus"
/proviral
/strain="type C"


```
/db_xref="taxon:61673"
/clone="43"
1. .629
467
LTR
misc_feature
primer_bind
632. .649
/feature="gly4"
1064. .2635
/feature="gag"
1064. .2635
/feature="gag"
/feature="gag"
/function="retroviral capsid"
/product="Gag protein"
join(1064. .2635,2639. .6223,6099. .8072)
/feature="viral polyprotein"
/codon_start=1
/protein_id="CAB65341.1"
/db_xref="GI:6688950"
/translation="MGQTVTTPSLTDHWEVRSRAHNSLVQVKKGPWQTFCASEWP
TFDQWPSSEGTNSEIILAVKAIIFQTGPSSHPDQEPYILWQDLAEDPPWPVKWPLN
KPKTPGPRILALCEKNKHSAAKEVSPRIYPIEIPPTWPEQVPVPPYPAQGAVRG
PSAPGAPVVEGPAAGTSSRRGATPERTDEIAILPLRTYGPMPGQGLQVWPFS
ADLYNWKTNHPPSEDPQRLTLGVLSELPESHQPTWDDCQQLQTLFTTEERERILLEA
RNVFCAGDRPQLQNEIDMGPELTPRGWDYNTAEGRESLKIYRALVAGUGASRRP
TNLAKVREVMQPNPEPSPVFLERLMEAFRRTPDPTSEAKQASVALAFIQSALDIR
KKLQLEGLQEALDLVREAEKYYRRETEEEKQREKEEREERDRRQKNTL
KILAAVEGKSSRERDRPKIRSGPROSGNLNRTPLDKQCAVCCKKGHWARNCKP
KGNKGPVLALBEDKDRGSDPLPEPRVALKVGQPVFLLVDTCAEHSVLLOPLGKL
KEKSKVMGATGQROYPWMTTRTVDLAGRVTHSELVPECPVPLGRDLITKMGQAI
SFEQGRPEVSNNKPIITVLTLQDDEYRLYSPQVQDDIQSWLEQFPQWAAETAGMG
LAKQVPOVILKASATPVSRQYPLSREAREGINPHVORLITQOGLVYVQSPNNTPL
LVRKPGNDYRPVQDLREVNKRQDIHTPTVPNYNLISALPPERNMYTVLDLDAF
CURLHPTSOPLFAFEWDPGTGRTGTLWTRLPQGFKNSTPFIDEALHRLDLNFRIOH
POVTLLOVDDLLLAGATKQDCLEGTALLLESLDLGYRASAKKQICRREVTYLGYS
LRGGRWTEARBKTVQIAPPTAKQVREFLTAGFCRLWIPGPAATLAAPLYLPIKE
KGEFWAHEQKTPDAIKKALLSAPALPDVTKFTLYLVDKRGVARGVLTQTLGPW
RPVAYLSKKDQPVASGMPVCLIAAVALIWKDADKLTGQNTIYIAPHALENIVRQ
PPDRWMTNARMTYQSLLTERTVTFAPALNPATLLPEETDEPVTDCHQLLIETG
VRKDLTDLPTGEVLTWFTDGSYYVEGKRWAGAAVVDGTHTWASSLPEGTSAQKAE
LMALTOALRAEGKSINITYDSRYAFATAHVGAHYKORGLLTSAGREIKNKEETLSL
LEALHLKRLAIHCPCGHQKADLISRGNOADRVAKAAQANLLPIETPKAPEPR
ROYTLEDQELIKKIDQFSETPGTCYTSYKREILPHKEGLEYVQOIHRLTHLGRKLO
QVRSYPHVULPVGADVSVWVHCVPQCLVNANPSRPPGKRLGSHPGHAEVDFTE
LPPAKYGNKLLVDFYDTSVWVEAYPTKETSTVVAKKILBEEIFRFGIPKIVGSDNG
PAFVAQVSGQKALIGLDWKLHCAYRPOSSQVVERMNTIKETILTKLTETGINDWIA
LLPFLVFRVNTPGQGLTPVELLYGPPPLVEIASVHADVLLSQPLFSLKALEWV
RORANKQIREAYSGEGDLQVPHRFQGVSVYVRHRAGNLETNRKGPYLVLVLTPTAV
KVEGLSTWISHAVLAPPPDSGWAERKTEMLKRLRLRLVPIYNNNSPQGMHPTLSW
RHLPTRGEPKRLRILPLSFASITANFLITTPQASSKRLIDSNNPHRPLSLTILIDP
DTGVTVNSTRVAPRGTWMPLEHFLRLINPAKSTPPNPLVRSYGFYCCPGTEKEKY
GGSESGFCRRSWCVTSNDGDKWPIISLODRYKFSFVNSGPGKYKVMKLYKOKSCSPD
LDYLKISFTEKGQENIOKINGMSWGVFVYKGGAGSTLITILRLIETGTEPPVAVG
PDKVLAEQGPALPPHNPVLPQTLSLRPDITQPPSNGTGLIPTNTPRNSPGVPVK
QORLPSLQGFAPALNSDTPDSCWLCSSGPPYEGMAKEGKFNVTKEHRNOCTW
SGRNKLTLETSVSGKGTCTGKAPPSSHLCYSTVVEQASENQYLPVGNRWACNTGL
TPCVSSVFNOSKFCVMQVIVPVRYHYPEEVLDEVDYRNRKPREPVSUTLAVMLG
LCTAVGCTGTAAALITPOOLEKGLGELHAAMTEDLRALESVSNLESLTSLSEVL
QNRKGLDILLFLREGGLCAALKECCFYVDHSGALRDSMSKLRELRERRRRREADQW
FKGWNRSPWMTLLSALTGLFLVLLLLTVGPCLINRFVAFVRERSAVQIMVLRQO
YQGLSQGETDL"
2639. .6223
/feature="pol"
/function="retroviral protease and polymerase"
/product="Pol protein"
2639. .6223
/feature="pol"
6099. .8069
/feature="env"
/function="retroviral envelope"
/product="Env protein"
6099. .8069
/feature="env"
8122. .8750
mat_peptide
gene
mat_peptide
gene
LTR
```

3339 GAGAGTCAATAAAAGGGTGCAGGACATACACCCAAACGGTCCCGAACCCCT 3388
251 TyrAsnLeuLeuSerAlaLeuProGluArgAsnTrpTyrThrValLe 267
3389 TATAACCTCTTGGCGCCCTCCCGCTGAACGGAACCTGGTACACAGTATT 3438
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284
3439 GGACTTAAAGATGCTTCTTCTGCTGAGATTACACCCCTAGCCCAAC 3488
284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
3489 CGCTTTTTCGCTTCAATGAGAGATCCAGGTACGGGAAGAACCGGGCAG 3538
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317
3539 CTCACCTGAGCCGACCTGCCCCAAGGGTTCAAGAACTCCCGACCATCTT 3588
317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
3589 TGACGAAGCCCTACACAGGACCTGCGCAACTTCAGGATCCAAACACCCCTC 3638
334 InValThrLeuLeuGlnTrpValAspLeuLeuAlaGlyAlaThr 350
3639 AGGTGACCCCTCTCCAGTACGTGGATGACCTTCTGCGGGGAGCCACC 3688
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAs 367
3689 AACAGGAGCTGCTTGAAGGTACAGAGGCATCTGCTGGAAATGCTCTGA 3738
367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgTrpGluV 384
3739 CCTAGGCTACAGACCTCTGCTTGAAGGCCAGATTTGCAGGAGAGAGG 3788
384 aThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400
3789 TAACATCTTGGGTACAGTTTGCGGGGCGGCGCAGCGATGGCTGACGGAG 3838
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG 417
3839 GCACGGGAAGAAACTGTAGTCCAGATACCGGGCCCCCAACACACGCCAAACA 3888
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
3889 AGTGAGAGAGTATTTGGGGACAGCTGGATTTTGCAGACTGTGGATCCCGG 3938
434 lypPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450
3939 GGTTCGGACCTTAGCAGCCCCCTCTACCCGCTAACCAAGAAAGAGG 3988
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467
3989 GAATTTCTCTGGCTCCTCAGCACCAGAGACATTTGATGCTATCAAAA 4038
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProaspValThrLysProP 484
4039 GGCCCTTGCTGAGCGCACCTGCTTGGCCCTCCTCAGCTTAACATAACCCCT 4088
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500
4089 TTACCCCTTTATGTGGATGAGCGTGAAGGAGTAGCCCGAGGAGTTTAAACC 4138
501 GlnThrLeuGlyProTrpArgProValAlaIaTyrLeuSerLysLysLe 517
4139 CAAACCCCTAGGACCATGGAGGACCTGTGCTACCTACCTGCTCAAGAAGCT 4188
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaAlav 534
4189 TGATCTGTAGCCAGTGGTGGCCCGATGCTGAAGGCTATCGCAGCTG 4238
534 aAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550
4239 TGCCCATCTGGTCAAGGAGCTGACAAATTCACITTTGGGACAGATATA 4288

551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567
4289 ACTGTATAGCCCCCATGATTTGGAGAACATCGTTGGCAGCCCCCAGA 4338
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuT 584
4339 CCGATGATGACCAACCGCCGATGACCCACTATCAAAAGCCTGCTTCTCA 4388
584 hrGluArgValThrPheAlaProAlaAlaLeuAsnProAlaThrLeu 600
4389 CAGAGGGGTCACTTTCGCTCCACCGCCGCTCTCAACCCCTGCCACTCTT 4438
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617
4439 CTGCTCAAGAGACTGATGAACCACTGACTCATGATTGCCATCACTATT 4488
617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634
4489 GATTGAGGAGACTGGGGTCCGCAAGGACCTTACAGACATACCGCTGACTG 4538
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650
4539 GAGAAGTGCCTAACCTGCTTCACTGACGGAAGCAGCTATGCTGGTGAAGGT 4588
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667
4589 AAGAGGATGCTGGGGCGCGGTGGTGACGGGACCCACACGATCTGGGC 4638
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684
4639 CAGCAGGCTGCCGGAAGAACTTTCAGCGCAAAAGGCTGAGCTCATGGCCC 4688
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700
4689 TCACCGAAGCTTTGGCGCTGGCCGAGGGAATCCATAAATTTATACG 4738
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717
4739 GACAGCAGTATGCTTTGGGACTGTCACAGCTACACGGGGCCATCTATAA 4788
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734
4789 ACAAGGGGGTGTCTTACCTCAGCGAGGAGGAATAAAGAACAAAGAGG 4838
734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750
4839 AAATTCAGGCTATTAGAACCTTACATTTGCCAAAAAGGCTAGCTATT 4888
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767
4889 ATACACTGTCTCGACATCAGAAAGCCCAAGATCTCATATCTAGAGGAA 4938
767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784
4939 CCAGATGGCTGACCGGGTGGCCACAGCAGGAGCCCGAGCTGTTAACCTTC 4988
784 euProfileIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800
4989 TGCCTATAATAGAAACCCCAAGCCCAAGAACCCAGAACCCAGACAGTACAC 5038
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817
5039 CTAGAAGACTGGCAAGAGATAAAAAAGATAGACCAGTTCTCTGAGACTCC 5088
817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834
5089 GGAGGGGACCTGCTATACCTCATATGGAGGAATCCTGCCCCCAACAG 5138
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850
5139 AAGGTTAGAATATGTCACAGATACATCGTCTCACCCACCTAGGAAC 5188

```
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867
|||||
5189 AAACACCTGCAGCAGTTGGTCAGAACATCCCTTATCATGCTTCTGAGGCT 5238
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884
|||||
5239 ACCAGGAGTGGCTGACTCGGTGGTCAAAACAATGTGTGCTCCCTGCCAGCTGG 5288
884 aAsnAlaAsnProSerArgIleProGlyLysArgLeuArgGlySer 900
|||||
5289 TTAATGCTAATCTTCAGAGTACTCCAGAAAGAGACTAAGGGGAGC 5338
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917
5339 CACCCAGGCGCTCACTGGGAGTGGACTTCACCTGAGGTAAAGCGGCTAA 5388
917 sTyrGlyAsnLysTrpLeuValPheValAspThrPheSerGlyTrpV 934
|||||
5389 ATACGGAACAATATCTATTGGTTTGTAGACACCTTTTCAGGATGGG 5438
934 alGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950
|||||
5439 TAGAGGCTTATCTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAAAA 5488
951 IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlyse 967
5489 ATACTGGAGAAATTTTCCAGATTGGAAATACCTTAAGTAATCGGGTC 5538
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984
|||||
5539 AGACAATGGTCCAGCTTTTTCGCCAGGTAACTCAGGAGCTGGCCAAAG 5588
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000
5589 TATTGGGATTGATGGAACTGCATTGCTACATACAGACCCCAAGACTCA 5638
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017
5639 GGACAGGTAGAGGATGAATAAACCATTAAAGAGACCTTACCAAAIT 5688
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034
|||||
5689 GACCAGAGACTGGCATTAAATGATTGGATAGCTCTCTCCCTTTGTGC 5738
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050
5739 TTTTATAGGTAGACACCCCTGGACAGTTGGCGCTGACCCCTATGAA 5788
1051 LeuLeuTyrGlyGlyProProLeuAlaGluIleAlaPheAlaHis 1067
5789 TTGCTCTACGGGGACCCCGCTGGTAGAAATTCCTTCGTACATAG 5838
1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084
5839 TGTCTAGTGTGCTTTCCAGCCTCTGTCTCTAGCTCAAGGGCGCTCG 5888
1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSerGly 1100
5889 AGTGGGTGAGGCACACGAGCTGGAACAGCTCCGGGAGCCCTACATCAGGA 5938
1101 ...GlyAspLeuGlnValProIleHisArgPheGlnValGlyAspSerValTy 1116
5939 GAAGGAGACTTGCAGCTTCCACATCGCTTCCAAGTGGGAGATTCAGTCTA 5988
1116 rValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyPro 1133
5989 TGTATAGCCCGCTGTCAGGAAACCTTCGAGACTCGGTGGAAGGCCCTT 6038
1133 yrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIlePro 1149
6039 ATCTCGTACTTTTGACCACACACGCGTGTGAAAGTCAAGGAATCTCC 6088
1150 ThrTrpIleHisAlaSerHisValLysProAlaProProAspSerG 1166
```

```
|||||
6089 ACCTGCATCATGATCCACCGTTAAGCTGGCGCCACCTCCCGACTCGGG 6138
1166 yTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgL 1183
|||||
6139 GTGGAGAGCCGAAAGACTGAGAAATCCCTTAAGCTTCGCTCCATCGCC 6188
1183 euValProTyrSerAsnAsnSerProGlyGln 1194
6189 TGGTTCTTACTCTAACAATAACTCCCGAGCCAG 6223
```

seq_name: gb_v1:PERY17013

seq_documentation_block:

LOCUS PERY17013 7808 bp DNA VRL 14-APR-2000
DEFINITION Porcine endogenous retrovirus gag, pol, env genes.

ACCESSION Y17013
VERSION Y17013.1 GI:3005985
KEYWORDS env gene; gag gene; pol gene.
SOURCE porcine endogenous retrovirus.
ORGANISM porcine endogenous retrovirus

Viruses; Retrovirdae; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.

REFERENCE 1 (bases 1 to 7808)
AUTHORS Czauderna,F., Fischer,N., Boller,K., Kurth,R. and Tonjes,R.R.
TITLE Establishment and characterization of molecular clones of porcine
endogenous retroviruses replicating on human cells

J. Virol. 74 (9), 4028-4038 (2000)
MEDLINE 20219394

REFERENCE 2 (bases 1 to 7808)

AUTHORS Toenjes,R.R.
TITLE Direct Submission

JOURNAL Submitted (30-MAR-1998) R.R. Toenjes, Paul-Ehrlich Institut,
Paul-Ehrlich Str. 51-59, P.O.Box 1740, D-63207 Langen, FRG

LOCATION/Qualifiers

FEATURES

source

gene

CDS

1..7808
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
/tissue_type="kidney"
/cell_line="PK15"
/gene="gag"
260..1834
260..1834
/gene="gag"
/codon_start=1
/protein_id="CAA76581.1"
/db_xref="GI:3005986"
/translation="MSQTVTTPLSLTLDHWTVEVRSRAHNSVQVKKGPMQTCASEWP
TEDVGPSEGTNSEIILAVKALIFQTGPSSHPDQEPYILTQDLAEDPPVWKPLN
KPKPGPRILALGKKNKSAEKVEPSRIYEIEPTWPEQPPVPPYPAQAVRG
PSAPPAGVVEGPAAGTSGRGATPERTDEIALLPTVYPPMPGQLOPLQWPFSS
ADLYNMKTNHPPTSEDPQLTGLVSLMF SHQPTWDCQQLQTLTFAERERILLEA
RKNVPGADGRPTQLQNEIDMGFPLTRLDWDYNTAEGTSLKIYRQALVAGLRGASRRP
TNLAKREVVGNGEPVSEVLERLMEAFRRFTPEPTSEAKASVALAFIGOSALDIR
KKLORLEGLQEAFLRDLVBEAKYVGRETEKEEQRKEEREEREERDRQENLT
KILAAVVEKSSRRERDRFRKIRSGPRQSGNLGNRTPLDKDQCAICYCKEKGHWARNCPK
KGNKGPVIALLEEDKD"

gene

CDS

2120..4726
/gene="pol"
/gene="pol"
/codon_start=1
/protein_id="CAA76582.1"
/db_xref="GI:3005987"
/translation="MCAQISFEQGRPEVSVNNKPTITLTQLDDEVRILYSPQVKPDQD
IQSWLEQFPQAAETAGMGLAKQVPPQVQLKASATPVSVROYPLSRAREGIPHWQ
RLIQGILYVQSPNTPPLPVKRGKTNDIRVQDLREVNKRQVDIHTVTPNPLLS
ALPPERNWTVIDLKDAPFLRLHPTSQPLFAFEWRDPTGQTGTLTRLPQGFKNR
PTIFDEALHRLANFRIOHPQVTLQYVDDLLAGATQDCLEGTAKLLILSDIYGR
ASAKKAQICREVTVLIGSLRGQRWLTARKKTVVQIPAPTTAKQVREFLTAGFCR
LWIPGFATLAAPLYPLTKEGEFSNAPEHQAFDAIKALLSAPALALPDVTKPTLY

VDRKRGVARGVLTQTLPGRWRPVAVLSKLLDPVAGSWPVCLKAI AAVAILVKDADKLT
 LGONITVIAHLENIVRQPPDRWMTNARMTHYQSLLLTERVTFAPFAALNPAFLPE
 ETDPEDVHOCHOLLIEETGRVKDLTDIPITGEVLTWFTDGSYVVEGRMAGAAVVDG
 THTIWASSLPESGSAKELMALTOALRLAEGKSINIYDSEYAFATAHVHCAIVKOR
 GLTISAGREIKNEELLSLELHLPLKRLAIITFCGHQKAKDLISRGQMADRVAKQA
 AQAVNLLPIETPRAPRPRQRTLEDWEIKKIDQFSETPEGTCTSYSGKEILLPHKEG
 LEYVOQIHLRTHLTKLQOLVRYTSFYHVRLEPVGADSVYKHCPCQVLNANPSRIPP
 GKRLRSGHPGAHWEVDETEVKPAKYNKYLLVFEVDTFSGWEAYPTKKTSTVVAKKY
 WRKFFRDLVLR"

5294..7267

/gene="env"

5294..7267

/gene="env"

/codon_start=1

/protein_id="CAA76583.1"

/db_xref="GI:3005988"

/db_xref="SPTREMBL:O73506"

/translation="MHPTLSWRHLPTRGGEKRLRIPLSFASIANPLILITITPQASSK
 RLIDNSNHRPLSLTWLIIDPDGTVTNSTRGVAGRTWMPHELFCRLINPAVKSTP
 PNLVRSYGFYCCPGTEBEKRYCGSEBSFCRWSCVTSNDGDKWFIISLDQRYKFSFN
 SGPGKYVMKLYDKCSPDLIDLYKISFTEKGQENIOKWIINGMSWGIYFYKYGGGA
 GPTLTIRLETGTETPPVAVGPDKVLAEQGPALPEPHNLPVOLTSLRPDITOPPSN
 GTGLIPTNTPRNSPGVPVTKGRLFSLIQAFQALNSTDPDATSCWLCILSSGPPY
 EGMAKSGKNVTKEHRNQTWSRNKLTITLTVSGKGTCTIGKAPPSHQHLCYSTVVEQ
 ASENQYLVPYNRWACNTGLTFCVSSVFNQSKDFVMVQIVPRVYHPEEVILDEY
 DYRNRPKREPVSLLTSLVMLGLTAVGVTGTAALITGPQOLEKGLGELHAAMTDLR
 ALBESVNLPELSLTSVVLQNRGLDLFLREGGLCAALKEECFCFYVDHSGAIRDS
 MSKLRELRERREREADOGWFEWNRSPWMTLLSALTGLPLVLLLLLTGVPOLIN
 RFVAFYRERVGAQVIMVLRQYQGLLSQGETDL"

BASE COUNT 2077 a 1971 c 2047 g 1713 t

ORIGIN

alignment_scores:

Quality: 6202.50 Length: 1195

Ratio: 5.217 Gaps: 1

Percent Similarity: 99.498 Percent Identity: 98.912

alignment_block:

US-09-171-553b-5 x PERV17013 ..

Align seg 1/1 to: PERV17013 from: 1 to: 7808

1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17
 |||||
 1835 GGGAGACGGGGTTCGACACCCCTCCCTGAGCCAGGGTAACTTTGAAGT 1884
 17 lGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34
 |||||
 1885 GGAGGGGCAACACAGTGTGAGTTCCTGGTATACCGGAGCGGAGCAATCAG 1934
 34 alLeuLeuGlnProLeuGlyLysLysGluLysLysSerTrpValMet 50
 |||||
 1935 TGCTGCTACACCATTAGGAAACTTAAAGAAAAAATCTGGTGGTGGTG 1984
 51 GlyAlaThrGlyGlnArgGlnTyrProTrpThrThrArgThrValAs 67
 |||||
 1985 GGTGCCACAGGCAACGGCAGTATCCATGGACTACCCGAAGAACCGTTGA 2034
 67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysP 84
 |||||
 2035 CTTGGCAGTGGGACGGGTAAACCATCGCTTCTGTGTCATCCCTGAGTGCC 2084
 84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100
 |||||
 2085 CAGTACCCCTCTTAGGTAGAGACTTACTGACCACAGATGGAGCTCAAT 2134
 101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117
 |||||
 2135 TCTTTTGAACAAGGAGACAGCAAGTGTCTGTGAATAACAACCCCATCAC 2184
 117 rValLeuThrLeuGlnLeuAspGluTyrArgLeuTyrSerProGlnV 134
 |||||

2185 TGTGTTGACCCCTCCAAATTAGATGATGAATATACGATATATTTCTCCCAAG 2234
 134 allLysProAspGlnAspIleGlnSerTrpLeuGlnPheProGlnAla 150
 |||||
 2235 TAAAGCCTGATCAAGATATACAGTCTCTGGTTGGAGCAGTTTCCCAAGCC 2284
 151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnVa 167
 |||||
 2285 TGGGCAAGAACCCGACGGGATGGTTTGGCAAGCAAGTTCCTCCCAAGGT 2334
 167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184
 |||||
 2335 TATTCAACTGAAGCCAGTGTCTACACAGTATCAGTCAGACAGTACCCCT 2384
 184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200
 |||||
 2385 TGAGTAGAGAGGCTCGAAGGAATTTGGCCGCATGTTCAAAGATTAAATC 2434
 201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217
 |||||
 2435 CAACAGGGCATCTAGTTCCTGCTCAATCCCTTGGAACTACTCCCTGCT 2484
 217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234
 |||||
 2485 ACCGTTTAGAAGCCCTGGGACCAATGATATTCACAGCTACAGGACTTGA 2534
 234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250
 |||||
 2535 GAGAGGTCAATAAAGGGTGCAGGACATACACCAACGGTCCCGAACCCCT 2584
 251 TyrAsnLeuLeuSerAlaLeuProGluArgAsnTrpTyrThrValle 267
 |||||
 2585 TATAACCTTGTAGCGCCCTCCCGCTGGAACGGAAGTGTGACACAGTATT 2634
 267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284
 |||||
 2635 GGACTTAAAGATGCTCTTCTCTGCTGAGATTACACCCACTAGCCAAC 2684
 284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
 |||||
 2685 CACTTTTTCCTTCCGAATGGAGAGATCCAGGTACGGGAAGAACCGGCGAG 2734
 301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317
 |||||
 2735 CTCACCTGGACCCGACTGCCCCCAAGGTTTCAAGAACTCCCGGACCATCTT 2784
 317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
 |||||
 2785 TGAGAAAGCCCTACACAGGACCTGGCCAACTTCAGGATCCAACACCCCTC 2834
 334 lnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350
 |||||
 2835 AGGTGACCCCTCTCCAGTAGTGATGACCTGCTTCTGGCGGGAGCCACC 2884
 351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuGluLeuSerAs 367
 |||||
 2885 AAACAGAGACTGCTTAGAAGGTACGAGGCACTACTGCTGGAATTTGCTGA 2934
 367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
 |||||
 2935 CCTAGGCTACAGACCTCTGCTAAGAGGCCAGATTTCAGGAGAGAGG 2984
 384 alThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400
 |||||
 2985 TAACATACTTGGGGTACAGTTTGGGGGGGGGCGGAGCGATGGCTGACGGAG 3034
 401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysGl 417
 |||||
 3035 GCACGGAAGAAACTGTAGTCCAGATACCGGCCCCCAACACAGCCARACA 3084
 417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
 |||||
 3085 AGTCAGAGAGTTTTGGGGGACAGCTGGATTTCGACAGACTGTGGATCCCG 3134

434 lypheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450
|||||
3135 GGTTCGACCTTAGCAGCCCACTCTACCCGCTAACCAAGAAAGGG 3184
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467
|||||
3185 GAATTCCTCTGGGCTCTGAGCACCAGAGGCAATTTGATGCTATCAAAA 3234
467 sAlaLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProP 484
|||||
3235 GGCCTCTGAGCGCACCTGCTTGGCCCTCCCTGACGTAACATAACCCCT 3284
484 heThrLeuTyrValAspGluArgLysGlyValAlaAlaArgGlyValLeuThr 500
|||||
3285 TTACCCCTTTATGTGATGAGGTAAGGAGTAGCCCGAGGAGTTTAAACC 3334
501 GlnThrLeuGlyProTrpArgProValAlaThrLeuSerLysLysLe 517
|||||
3335 CAACCCCTAGACCATGGAGGAGACCTGTTCCTACCTGTCAAGAGAGCT 3384
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaAlay 534
|||||
3385 TGATCTCTGACCGAGTGTGGCCGCTATGCCCTGAAGGCTATCGCAGCTG 3434
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550
|||||
3435 TGGCCATACATGGTCAAGSACGCTGACAAATTGACCTTGGGACAGATATA 3484
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567
|||||
3485 ACTGTAATAGCCCCCATGATGATTTGGAGAACATGTTTCGGCAGCCCCAGA 3534
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuT 584
|||||
3535 CCGATGATGACCAACCCCGCATGACCCACTATCAAGGCTGCTCTCA 3584
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600
|||||
3585 CAGAGAGGGTCACCTTCCTGCCACCAGCCGCTCTCAACCCCTGCCACTCT 3634
601 LeuProGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617
|||||
3635 CTGCTGAAGAGACTGATGAACCAAGTGCATCATTTGCCATCAACTATT 3684
617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634
|||||
3685 GATTGAGGAGACTGGGCTCCGCAAGGACCTTACAGACATACCCGCTGACTG 3734
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650
|||||
3735 GAGAAGTCTAACTGGTTCACTCAGCGAAGCAGCTATGTGGTGAAGGT 3784
651 LysArgMetAlaGlyAlaAlaValAlaValAspGlyThrArgThrIleTrpAl 667
|||||
3785 AAGAGGATGCTGGGGCGGGGTGGTGGACGGGACCCACACGATCTGGGC 3834
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684
|||||
3835 CAGCAGCCTCGCGAAGAACTTACGCGCAAAAGGCTGAGCTCATGGCCC 3884
684 euthrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700
|||||
3885 TCACGCAAGCTTTCGCGCTGCGCGAAGGAAATCCATAACATTTATACG 3934
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717
|||||
3935 GACAGCAGGATGCTTGTGGACTGTCACACGTACACGGGGCCATCTATAA 3984
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734
|||||
3985 ACAAGGGGTTGCTTACCTCAGCAGGGGGGAAATAAGAACAAAGAGG 4034

734 luLeuLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750
|||||
4035 AAATTCCTAAGCCTATTAGAAGCCTTACATTTGCCAAAAGGCTAGTATT 4084
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767
|||||
4085 ATACACTGTCTTGGACATCAGAAAGCCAAAGATCTCATATCTAGAGGAA 4134
767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784
|||||
4135 CCAGATGGCTGACCGGGTTGCCAAGCAGCAGCCAGGCTGTAACTTTC 4184
784 euProIleLeuGluThrProLysAlaProGluProArgArgGlnTyrThr 800
|||||
4185 TGCCTTATATAGAAACGCCAGAGCCCAAGAACCCAGACAGACAGTACACC 4234
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817
|||||
4235 CTAGAAGACTGGCAAGAGATAAAAGATAGACCAGTCTCTGAGACTCC 4284
817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834
|||||
4285 GAGGGGACCTGCTATACCTCATATGGGAAGAAATCTGCCCCACAAAG 4334
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850
|||||
4335 AAGGGTTAGAAATATGTCCAACAGATACATCGTCTCACCCACCTAGGAAC 4384
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867
|||||
4385 AAACACCTGCAGCAGTTGGTCAGAACATCCCTTATCATTTCTGAGGCT 4434
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884
|||||
4435 ACCAGAGTGGCTGACTCGGTGTCAACATTTGTGTGCTGCCAGCTGG 4484
884 alAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900
|||||
4485 TTAATGCTAATCTTCCAGAAATACCTCCAGGAAGAGACTAAGGGGAAGC 4534
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917
|||||
4535 CACCCAGCGCTCCTGGGAAGTGGACTTCACTGAGGTAAAGCCGGCTAA 4584
917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934
|||||
4585 ATACGGAACAAATATCTATTGGTTTTGTAGACACCTTTTCAGGATGG 4634
934 alCluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950
|||||
4635 TAGAGGCTTATCCTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAAAA 4684
951 IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967
|||||
4685 TA CTGGAGGAAATTTTCCGAGATTTGGAAATACCTAAGGTAAATCGGGTC 4733
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984
|||||
4734 ACACAAATGGTCCAGCTTTTGTGGCCAGGTAAGTACAGGACTGGCCAAGA 4783
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000
|||||
4784 TATTGGGATTTGATTGGAACCTGCATTTGTGCATACAGACCCCAAGCTCA 4833
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017
|||||
4834 GGACAGGTAGAGAGGATGAATAGAACCATTAAAGAGACCTTTACCAAAAT 4883
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034
|||||
4884 GACCACAGAGACTGGCATTAAATGATGGATAGTCTCTCTGCTTTGTGC 4933
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050

LINRPFVPRBQVSQVAVRIMWLRLQQVQGLPS"
 ITR 8216..8918
 polyA_signal 8810..8815
 BASE COUNT 2435 a 2174 c 2253 g 2056 t
 ORIGIN

alignment_scores:
 Quality: 6142.00 Length: 1191
 Ratio: 5.196 Gaps: 0
 Percent Similarity: 99.244 Percent Identity: 97.565

alignment_block:

US-09-171-553B-5 x PEN293656 ..

Align seg 1/1 to: PEN293656 from: 1 to: 8918

1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuIysVa 17
 2728 GGAAGACGGGTTCCGACCCCTCCCGAGCCCGAGGGTAACCTTTAAAGGT 2777
 17 lGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34
 2778 GGAGGGCAACCAAGTTGAGTTCTGTTGATCCGCGAAGAACATTCAG 2827
 34 allLeuLeuGlnProLeuGlyLysLeuLysGlyLysLysSerTrpValMet 50
 2828 TGCTACTACAGCCATTAGGAAACTAAAGATAAAAAATCCTGGGTGATG 2877
 51 GlyAlaThrGlyGlnArgGlnTyrProTrpThrArgArgThrValAs 67
 2878 GGTGCCACAGGGCAACAACAATATCCATGGACTACCCGAAGAAGACAGTTGA 2927
 67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysP 84
 2928 CTTGGGAGTGGAGGGTAACCCACTCGTTCTGGTCATACCTGAGTGCC 2977
 84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100
 2978 CAGCACCCCTCTTAGGTAGAGACTTACTGACCAAGATGGGAGCACAAAT 3027
 101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117
 3028 TCCTTTGAACAAGGGAACAGAGTGTCTGCAAAATAACAAACCTATCAC 3077
 117 rValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnV 134
 3078 TGTGTTGACCTCCCAATTAGATGACGAATATCGACTATATTCCTCCCTAG 3127
 134 allYsProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150
 3128 TAAAGCCTGATCAAAATATACAAATCTGTTGGAACAGTTTCCCAAGCC 3177
 151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnVa 167
 3178 TGGCAGAAACCGCAGGGATGGGTTTGGCAAGCAAGTTCCCCACAGGT 3227
 167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184
 3228 TATTCACTGAAGCCAGTGTCTGCACCAAGTGTCACTCAGACAGTACCCCT 3277
 184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200
 3278 TGAGTAAAGAAGCTCGAAGGAATTCGGCGGCATGTTCAAAAGATTAAATC 3327
 201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217
 3328 CAACAGGCGATCCTAGTTCCTGTCCCAATCTCCCTGGAATACCTCCCTGCT 3377
 217 uproValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234
 3378 ACCGGTTAGAAAGCCCTGGGACTAATGACTATATCGACCAAGTACAGCACTTGA 3427

234 rGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250
 3428 GAGAGTCAATAAACGGGTGCAGGATATACACCAACAGTCCCGAACCT 3477
 251 TyrAsnLeuLeuSerAlaLeuProGluArgAsnTrpTyrThrValLe 267
 3478 TATAACCTCTTGTGTCTCCACCCCAACGGAGCTGGTATACAGTATT 3527
 267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284
 3528 GGACTTAAAGGATGCTTCTTCGCTGAGACTACATCCCACTAGCCAAAC 3577
 284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
 3578 CACTTTTTCCTTCGAATGGAGATCCAGGTGCGGGAAGAACCCGGGAG 3627
 301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317
 3628 CTCACCTGGACTCGACTGCCCAAGGGTTCAAAAACCTCCCGACCATCTT 3677
 317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
 3678 TGACGAGGCCCTACACAGAGACTGGCCAACTTCAGGATCCCAACACCCC 3727
 334 lNValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350
 3728 AGGTGACCTCTCCAGTACGTGGATGACCTGCTTCTGGCGGAGGCCACC 3777
 351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAs 367
 3778 AAACAGGACTGCTTAGAAGTACGAGGACACTACTGCTGGAATTTGTGGA 3827
 367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
 3828 CTTAGCTACAGAGCTCCGCTAAGAAGGCCAGATTTGCAGAGAGAGG 3877
 384 alThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400
 3878 TAACATACTTGGGTACAGTTTTCGGGGCGGCGAGCGATGGCTGACGGAG 3927
 401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG 417
 3928 GCACGGAAGAGAACTGTAGTCCAGATACCGGCCCAACCCACAGCCAAACA 3977
 417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
 3978 AGTGAGAGAGTTTTTCGGGACAGCTGGATTTTCGACACTGTGGATCCCGG 4027
 434 lypPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGlyLysGly 450
 4028 GGTTCGACCTTAGCAGCCCTCTACCTACCCATAACCAAGAAAGAGG 4077
 451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467
 4078 GAATTCCTCTGGCTCCTGAGCACCAAGAGGCAATTTGATGCTATCAAAA 4127
 467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProP 484
 4128 GGCCCTGCTGAGCCCTCTCTGCGCCCTCCCTGACGTGACTAAACCCCT 4177
 484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500
 4178 TTACCTTTTATGTGGATGAGCTGAAGGAGTAGCCCGGGAGTTTAAACC 4227
 501 GlnThrLeuGlyProTrpArgProValAlaTyrLeuSerLysLysLe 517
 4228 CAACCTTAGACCATGGAGAGACCTGTTCCTACCTGTCRAAAGAGCT 4277
 517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaAlaV 534
 4278 CGATCTCTGAGCCAGTGGTTGGCCCATATGCTGGAAGGCTATCGCAGCTG 4327
 534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550

||||| 4328 TGCCATACCTGGTCAAGGAGCGCTGACAAATTTGACTTTGGACAGAATATA 4377
||||| 551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567
||||| 4378 ACTGTAATAGCCCCCATCGGTTGGAGACATCGTTCCGACGCCCCAGA 4427
567 pArgTTPMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuT 584
||||| 4428 CCGATGGATGACCAACGCCCGCATGCCACATATCAAAAGCTGTCTCA 4477
584 hrcGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600
||||| 4478 CAGAGAGGTCACGTTCCGCTCCACCGAGCGCTCTCAACCTGCCACTCTT 4527
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617
||||| 4528 CTGCCTGAAGAGACTGATGACCACTGACTCATGATGTCATCACTATT 4577
617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634
4578 GATTGAGGAGACTGGGCTCGGAAGACCTTACAGACATACCGCTGACTG 4627
634 lYGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650
||||| 4628 GAGAAATGTTAACCTGGTTCACTGACGGAAGCAGCTATATGTGGAAGGT 4677
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667
4678 AAGAGGATGGCTGGGCGCGCGTGTGGACGGGACCGCCGACGATCTGGGC 4727
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684
4728 CAGCAGCCTGCGGGAAGAACTTACGACAAAGAGCTGAGCTCATGGCCC 4777
684 euThrClnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700
4778 TCACGCAAGCTTTGGCGCTGGCGGGAAGGAAATCCATAAATTTATACA 4827
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717
4828 GACAGCAGGTATGCTTTGGCAGCTGCACACGTACACGGGGCCATCTATAA 4877
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734
4878 GCAAAGGGGGTGTCTTACTCTACGAGGGAGGGAATAAAGAACAAAGAGG 4927
734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750
4928 AAATTTAAGCCTATTAGAGCCTTACATTTGCCAAAAGGCTAGCTATT 4977
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767
4978 ATACACTGTCTTGGACATCAGAAAGCCAAAGATCCATATCCAGAGGAA 5027
767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784
5028 CCAGATGGCTGACCGGGTGGCCAAAGCAGGAGCGCCAGGGTGTAAACCTTC 5077
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800
5078 TGCCCTATGATGAACACCCAAAGGCCCAAGACCCGCGACGACGATACACC 5127
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817
5128 CTAGAAGACTGCGAAGAAATAAAAAAGATAGACCAGTTCTCTGAAACTCC 5177
817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834
5178 GGAGGGGACCTGCTATACCTCAGATGGGAAGAAATCCTGCCCCACAAAG 5227
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850
|||||

5228 AAGGTTTAGAATATGTCACACAGATACATCGCTTAACCCACCTAGGAAC 5277
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867
5278 AAACACCTGCAGCAGCTTGGTCAGAACATCCCTTATCATGTTCTGAGGCT 5327
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884
5328 ACCAGGAGTGCTGACTCGGTGGTCAAAACATTTGTGCCCTGCCAGCTGG 5377
884 alAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900
5378 TTAATGCTAATCTCTCCAGAAATACCTCCAGGAAAGAGACTAAGGGGAGC 5427
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917
5428 CACCCAGGCGCTCACTGGGAAGTGGACTTCACTGAGGTAAAGCCGGCTAA 5477
917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934
5478 ATACGGAACAAATATCTATTGTTTGTGTAGACACCTTTTCAGGATGGG 5527
934 alGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950
5528 TAGAGGCTTATCTTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAGAAA 5577
951 IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967
5578 ATACTGGAGGAAATTTTCCAAAGATTTGGAAATACCTAAGGTAAATAGGTC 5627
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984
5628 AGACAAATGGTCCAGCTTTCTGTCGCCAGGTAAAGTCAGGGACTGCCCAAG 5677
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000
5678 TATTGGGGATTTGATTGGAAATCGCATTTGTGCATACAGACCCCAAGCTCA 5727
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017
5728 GGACAGGTAGAGAGGATGAATAGAACCTTTAAGAGACCCCTTACCAAT 5777
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034
5778 GACCACAGAGACTGGCATTAAATGATTGGATGGCTCTCCTGCCCTTTGTGC 5827
1034 euPheArgValArgAspThrProGlyGlnPheGlyLeuThrProTyrGlu 1050
5828 TTTTGGGGTGAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATGAA 5877
1051 LeuLeuTyrGlyGlyProProProLeuAlaGluIleAlaPheAlaHisSe 1067
5878 TTGCTCTACGGGGGACCCCGCTGGCAGAAATTTGCCTTTGCACATAG 5927
1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084
5928 TGCTGATGTGCTGCTTCCAGCGCTTTGTCTCTAGGCTCAAGCGCTCG 5977
1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSerGly 1100
5978 AGTGGGTGAGGCGAGCGGTGGAAGCAGCTCCGGGAGGCGCTACTCAGGA 6027
1101 GlyAspLeuGlnValProHisArgPheGlnValIleGlyAspSerValTyrVa 1117
6028 GGAGACTTGCAAGTTCCACATCGCTTCCAAAGTTGGAGATTCAGTCTATGT 6077
1117 lArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProTyrL 1134
6078 TAGACGCCACCGTGCAGGAAACCTCGAGACTCGGTGGAAAGGACCTTATC 6127
1134 euValLeuLeuThrThrProThrAlaValLysValGluGlyIleProThr 1150
6128 TCGTACTTTTGACCAACCAACGCGTGTGAAGTCGAAGAAATCCCCACC 6177

1151 TrpIleHisAlaSerHisValLysProAlaProProAspSerGlyTr 1167
 1167 TGGATCCATGATCCACAGTTAAGCGCGCCACCTCCCGATTCGGGGTG 6277
 1167 pArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgLeuV 1184
 6228 GAAAGCGGAAAGACTGAAATCCCTTAAAGCTTCGCTCCATCGCGGTG 6277
 1184 alProTyrSerAsnAsnSer 1191
 6278 TTCCTTACTCTGTCATAACTCC 6300

seq_name: gb_pat:ARI130475

seq_documentation_block:
 LOCUS ARI130475 8132 bp DNA PAT 16-MAY-2001
 DEFINITION Sequence 3 from patent US 6190861.
 ACCESSION ARI130475
 VERSION ARI130475.1 GI:14118800

KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 UNCLASSIFIED.

REFERENCE 1 (bases 1 to 8132)

AUTHORS Fishman,J.A.

TITLE Molecular sequences of swine retroviruses method of using
 JOURNAL Patent: US 6190861-A 3 20-FEB-2001;

FEATURES
 Location/Qualifiers
 source
 1..8132
 /organism="unknown"

BASE COUNT 2248 a 1977 c 2037 g 1870 t
 ORIGIN

alignment_scores:

Quality: 6099.50 Length: 1191
 Ratio: 5.173 Gaps: 1
 Percent Similarity: 98.992 Percent Identity: 96.893

alignment_block:

US-09-171-553b-5 x ARI130475 ..

Align seg 1/1 to: ARI130475 from: 1 to: 8132

1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17
 1167 TGGATCCATGATCCACAGTTAAGCGCGCCACCTCCCGATTCGGGGTG 6277
 2160 GGGAGACGGGTTCGGACCCCTCCCGAGCCCGAGGGTAACCTTGAAGGT 2209
 17 lGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34
 2210 GGAGGGCAACCAAGTTGAGTTCCTGTTGATACCGGAGCGCAACATTCAG 2259
 34 allLeuLeuGlnProLeuGlyLysLeuLysLysLysSerTrpValMet 50
 2260 TGCTACTACAGCCATTAGGAAACTAAAGATAAAAAATCCTGGGTGATG 2309
 51 GlyAlaThrGlyGlnArgGlnTyrProTyrThrArgThrValAs 67
 2310 GGTGCCACAGGGCAACACAGTATCCATGGACTACCCGAAGACAGTTGA 2359
 67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysp 84
 2360 CTTGGGAGTGGGAGGGGTAAACCCACACGTTCTTGGTCATACCTAGTGCC 2409
 84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100
 2410 CAGCACCCCTCTTAGGTAGAGACTTATTGACCAAGATGGGAGCACAAAT 2459
 101 SerPheGluGlnGlyArgProGluValSerValAsnLysProIleth 117
 2460 TCTTTTGAACAGGGAACAGAGAGTGTCTGCAATAATAACAACCTATCAC 2509

117 rValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnV 134
 2510 TGTGTTGACCCCTCCCAATTAGATGACGAATATCGACTATACTCTCCCTAG 2559
 134 allYsProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150
 2560 TAAAGCTGATCAAAATATACAATCTTGTGTTGGAACAGTTTCCCAAGCC 2609
 151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProGlnVa 167
 2610 TGGGCAGAAACCGCAGGATGGTTTGGCAAGCAAGTTCCCCCAAGT 2659
 167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184
 2660 TATTCAACTGAAGCCAGTCCACACCACTGTCAGTCAGACAGTAGTACCCT 2709
 184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200
 2710 TGAGTAAAGAAGCTCAAGAAGGAATTCGGCCGCATGTCCAAAGATTAATC 2759
 201 GlnGlnGlyIleLeuValProValGlnSerProTyrAsnThrProLeuLe 217
 2760 CACACAGGCATCTAGTCTCTGTCATCTCCCTGGAATACTCCCTGCT 2809
 217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234
 2810 ACCGGTTAGAAAGCCTGGGACTTAATGACTATTCGACAGTACAGGACTTGA 2859
 234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250
 2860 GAGAGGTCAATAAAGCGGTGCGAGGATATACACCAACAGTCCCGAACCT 2909
 251 TyrAsnLeuLeuSerAlaLeuProGluArgAsnTrpTyrThrValle 267
 2910 TATAACCTCTTGTGCTCTCCACCCCAAGGTTCAAGAACTCCCCGACCATTT 2959
 267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284
 2960 GGACITAAAGGATGCTTCTTCTGCTGAGATTACACCCCACTAGCCAC 3009
 284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
 3010 CACTTTTTCCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCGGCGAG 3059
 301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIleph 317
 3060 CTCACCTGGACCCGACTGCCCAAGGTTCAAGAACTCCCCGACCATTT 3109
 317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
 3110 TGACGAAGCCCTACACAGAGACCTGGCCAACTTCAGGATCCCAACACCTC 3159
 334 InValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350
 3160 AGGTGACCTCTCCAGTACGTGGATGACCTGCTTCTGGCGGAGGCCACC 3209
 351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAs 367
 3210 AACAGGACTGCTTAGAGGCGACGAGGCACCTACTGCTGGAATTTGCTGA 3259
 367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
 3260 CTTAGGCTACAGAGCTCTGCTTAAGAGCCCGCAGATTTGCAGAGAGAGG 3309
 384 alThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400
 3310 TAACATACTTGGGTACAGTTTGGCGGACGGCGAGCGATGGCTGACGGAG 3359
 401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG 417
 3360 GCACGGGAAGAAACTGTAGTCCAGATACCGGCCCAACACACAGCCAAACA 3409
 417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434

|||||:::||||| 3410 AATGAGAGAGTTTGGGACAGCTGGATTTTGCAGAGTGTGGATCCCGG 3459
434 lypheAlaThrLeuAlaProLeuTyrProLeuThrLysGluLysGly 450
3460 GGTTCGACACCTTACAGCCACCTCTACCCGCTAACCAAGAAAAGGG 3509
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467
3510 GAATTCCTCTCGGGCTCCTGAGCACCAGAGGCAATTTGATGCTATCAAAA 3559
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484
3560 GGCCCTGTGAGCGACAGCTGCTTGGCCCTCCCTGACGTAACTAAACCCCT 3609
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500
3610 TTACCCCTTTATGTGATGAGCGTAGGGAGTAGCCCGGGAGTTTAAACC 3659
501 GlnThrLeuGlyProTrpArgProValAlaTyrLeuSerLysLysLe 517
3660 CAAACCCCTAGGACCATGAGAGACCTGTCGCCCTACCTGTCAAAAGAGCT 3709
517 uaspproValAlaSerGlyTrpProValCysLeuLysAlaIleAlaIy 534
3710 CGATCCTGTAGCCAGTGTGTGGCCCATATGCTGAAGGCTATCGCAGCTG 3759
534 aAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyClnAsnIle 550
3760 TGGCCATACTGGTCAAGGACCTGACAAATTTGACTTTGGGACAGAAATA 3809
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProAs 567
3810 ACTGTAATAGCCCCCATGTCATTTGGAGAACATCGTTGCGCAGCCCCAGA 3859
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrClnSerLeuLeuLeu 584
3860 CCGATGGATGACCAACGCCGCGTACCCACTATCAAAAGCCTGCTCTCA 3909
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600
3910 CAGAGAGGTCAGCTTCGCTCCACGAGCGCTCTCAACCCCTGCCACTCTT 3959
601 LeuProGluGluThrAspGluProValThrHisaspCysHisGlnLeuLe 617
3960 TTGCTGAAGAGACTGATGAACACAGTGACTCATGATGCCATCAACTATT 4009
617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634
4010 GATTGAGGAGACTGGGGTCCGCAAGGACCTTACAGACATACCGCTGACTG 4059
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValIleGly 650
4060 GAGAAAGTCTAACCTGGTTCTACTGACGGAAGCAGCTATGTGGTGAAGGT 4109
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667
4110 AAGAGGATGGTGGGGCGCGGTGGTGGACGGGACCGCCGACAGTCTGGGC 4159
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684
4160 CAGCAGCCTGCGGAAGAACTTACAGCAAAAGGCTGAGCTCATGGCCC 4209
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700
4210 TCACGCAAGCTTGGGGCTGGCCGAGGCAATCCATAAATTTATACG 4259
701 aspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717
4260 GACACAGGTATGCTTTGGCAGCTGCACACGTACATGGGGCCATCTATAA 4309
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734
|||||:::||||| 4310 ACAAGGGGCTGCTTACCTACGAGGGAGGAAATAAGAACAAAGAGG 4359
734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750
4360 AATTTCTAAGCCCTATTAGAAGCGGTACATTTACCAAAAAGGCTAGCTATT 4409
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767
4410 ATACACTGTCTGGACATCAGAAAGCTAAAGATCTCATATCCAGAGGAAA 4459
767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784
4460 CCAGATGGCTGACCGGTTGCCAAGCAGGAGCCCGAGGGTGTAAAGCTTC 4509
784 euProIleLeuGluThrProLysAlaProGluProArgArgGlnTyrThr 800
4510 TGCCCTAATAGAAATGCCAAAGCCCAAGACCCAGACGACGACATACACC 4559
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817
4560 CTAGAAAGCTGGCAAGAGATAAAAAAGATAGACCACTTCTCTGAGACTCC 4609
817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834
4610 GSAAGGACCTGCTATACCTCAGATGGGAAGAAATCCTGCCCCACAAAG 4659
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850
4660 AAGGTTAGAAATATGTCACACAGATACATCGTCTAACCCACCTAGGAAC 4709
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867
4710 AAACACCTGCAGCAGTGTGTCAGAACATCCCTTATCATGTTCTGAGGCT 4759
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884
4760 ACCAGAGTGGCTGACTCGGTGTCAACATTTGTGTGCCCTGCCAGCTGG 4809
884 aAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900
4810 TTAATGCTAATCTCTCAGAAATGCCCTCCAGGGAAGAGACTAAGGGGAAG 4859
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917
4860 CACCCAGGCGCTCCTGGGAAGTGGACTTCACTGAGGTAAAGCCGCTAA 4909
917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934
4910 ATACGGAACAATACCTATTGCTTTTGTAGACACCTTTTCAGGATGG 4959
934 alGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950
4960 TAGAGGCTTATCCTACTAAGAAAGAGACTTCAACCCGTGGTGGCTAAAAA 5009
951 IleLeuGluIlePheProArgPheGlyIleProLysValIleGlySe 967
5010 ATACTGGAAGAAATTTTCCAGATTGGAATACCTTAAGGTAAATAGGGTC 5059
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984
5060 AGACAATGGTCCAGCTTTGTCGCCAGGTAAAGTACGAGGACTGGCCCAAG 5109
984 leuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000
5110 TATTTGGGATTTGATTTGAACTGCTATTTGTCATACAGACCCCAAGCTCA 5159
1001 GlyGluValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017
5160 GGACAGGTAGAGAGATGAATAGACCATTAAGAGACCCCTTACTAAAT 5209
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValI 1034
5210 GACCCGGAGAGCTGGCGTTAATGATTGGATAGCTCTCTGCCCTTTGTGC 5259

```

1034 euPheArgValArgAsnThrProGlyClnPheGlyLeuThrProTyrGlu 1050
|||||
5260 TTTTAGGGTTAGAAACACCCCTGGACAGTTTGGGCTGACCCCTATGAA 5309
|||||
1051 LeuLeuTyrGlyGlyProProLeuAlaGluIleAlaPheAlaHisSe 1067
|||||
5310 TTACTCTACGGGGACCCCCCATTTGGTAGAAATTCCTTCTGTACATAG 5359
|||||
1067 rAlaSpValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084
|||||
5360 TGCTGACGTCTGCTTCCAGCCTTTGTCTCTAGGCTCAAGGCACITG 5409
|||||
1084 LuTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSer... 1099
|||||
5410 AGTGGGTGAGACAAACGAGCGTGGAGGCAACTCCGGGAGGCCTACTCAGGA 5459
|||||
1100 GlyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValty 1116
|||||
5460 GGAGGAGACTTGCAGATCCACATCGTTTCCAAAGTGGGAGATTCAGTCTA 5509
|||||
1116 rValArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProT 1133
|||||
5510 CGTTAGACGCCACCGTCCAGGAAACCTCGAGAGTCGGTGGAGGGCCCTT 5559
|||||
1133 yrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIlePro 1149
|||||
5560 ATCTCTACTTTTGGACCAACCAACGCGTGTGAAGTCCGAAGGAATCTCC 5609
|||||
1150 ThrTrpIleHisAlaSerHisValLysProAlaProProAspSerG1 1166
|||||
5610 ACCTGGATCATGCATCCACCGTTAAACCGCGGCCACCTCCCGATTCGGG 5659
|||||
1166 yTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgL 1183
|||||
5660 GTGGAAGCGGAAAGACACTGAAATCCCTTAAGCTTCGCCCTCCATCGCG 5709
|||||
1183 euValProTyrSerAsnAsn 1190
:::|||||
5710 TGGTTCTTACTCTGTCAATAAC 5732

seq_name: gb_pat:AX052635

seq_documentation_block:
LOCUS AX052635 4402 bp DNA PAT 12-JAN-2001
DEFINITION Sequence 20 from Patent WO0071726.
ACCESSION AX052635
VERSION AX052635.1 GI:12226825
KEYWORDS
SOURCE
ORGANISM
porcine endogenous retrovirus.
retroviruses; 1-Mammalian type C virus group.
1 (bases 1 to 4402)
Fedorpiel,M.J.
Methods to inhibit infectious agent transmission during
xenotransplantation
Patent: WO 0071726-A 20 30-NOV-2000;
Mayo Medical Ventures (US)
FEATURES
Location/Qualifiers
source
1..4402
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
BASE COUNT 1259 a 1085 c 1111 g 947 t
ORIGIN

alignment_scores:
Quality: 6095.50 Length: 1192
Ratio: 5.174 Gaps: 1
Percent Similarity: 98.826 Percent Identity: 96.980

alignment_block:
1 GlyArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17
|||||
224 GGAAGACGGGGTTCGACCCCTCCCGAGCCAGGGTAACCTTTGAAGGT 273
|||||
17 IGLuGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34
|||||
274 GGAGGGCAACACAGTTGAGTTCTCTGGTTGATACCGAGCGGAACATTCAG 323
|||||
34 aLeuLeuGlnProLeuGlyLysLeuLysGluLysLysSerTrpValMet 50
|||||
324 TGCTACTACAGCCATTAGGAAACTTAAAGATATAAAAAATCTCTGGGTGATG 373
|||||
51 GlyAlaThrGlyGlnArgGlnTyrProTrpThrArgArgThrValAs 67
|||||
374 GGTGCCAGGGCAACACAGTATCCATGGACTACCCGGAAGAACAGTTTGA 423
|||||
67 pleuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysP 84
|||||
424 CTTGGAGTGGGACGGGTAAACCCACTCGTTTCTGGTCATACCTGAGTGCC 473
|||||
84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100
|||||
474 CAGCACCCCTCTTAGGTAGAGACTTATTGACCAAGATGGGAGCACAAATT 523
|||||
101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117
|||||
524 TCTTTTGAACAAAGGAAACACAGAGTGTCTGCAATATACAAACCTATCAC 573
|||||
117 rValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnV 134
|||||
574 TGTGTTGACCTCCCAATTAGATGAGATATATCGACTATPACTCTCCCTAG 623
|||||
134 alLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150
|||||
624 TAAAGCCTGATCAAAATATACAATTCGTGGTTGGAACAGTTTCCCAAGCC 673
|||||
151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProGlnVa 167
|||||
674 TGGGCAGAAACCGCAGGGATGGGTTTGGCAAGCAAGTTCCCCCAACAGT 723
|||||
167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTrpProL 184
|||||
724 TATCAACTGAAGCCAGTCCACACAGTGTCTAGTCAGACAGTACCCCT 773
|||||
184 euSerArgGluAlaArgGlyIleTrpProHisValGlnArgLeuIle 200
|||||
774 TGAGTAAAGAAAGCTCAAGAAGAAATTCGGCGCATGTCCAAAGATTAATC 823
|||||
201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217
|||||
824 CAACAGGCGCATCTAGTTCCTCTCCCAATCTCCCTGGAAATACTCCCTGCT 873
|||||
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234
|||||
874 ACCGGTTAGAAAGCCTGGGACTAATGACTATCGACAGTACAGGACTTGA 923
|||||
234 rGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250
|||||
924 GAGAGGTCAATAAACGGGTGACGATATACACCCCAACAGTCCCGAACCT 973
|||||
251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValLe 267
|||||
974 TATAACCTCTGTCTCTCCACCCCAACGAGCTGGTATACAGTATT 1023
|||||
267 uAspLeuLysAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284
|||||
1024 GGACTTAAAGGATGCCCTTTTCTGCCCTGAGATTACACCCACCTAGCCAAC 1073
|||||
284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyThrGlyGln 300
|||||

```

us-09-171-553b-5 x AX052635

Align seg 1/1 to: AX052635 from: 1 to: 4402

||||| 1074 CACTTTTTCCTTCCAAATGGAGATCCAGGTACGGGAAGAACCGGCGAG 1123
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePhe 317
1124 CTCACCTGGACCGGACTGCCCAAGGGTTCAAGAACTCCCGACCACTTT 1173
317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
1174 TGACCAAGCCCTACACAGAGACCTGGCCCACTTCAGGATCCACACCCCTC 1223
334 InValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350
1224 AGGTGACCCCTCCTCAGTACGTGGATGACCTGCTTCTGGCGGAGCCACC 1273
351 LysGlnAspCysLeuGlnGlyThrLysAlaLeuLeuLeuLeuSerAs 367
1274 AAACAGGACTGCTTAAAGGACGGAAGGCACTACTGCTGGATGTCTGA 1323
367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
1324 CCTAGGCTACAGAGCCTCTGCTAAGAAGGCCAGATTTGCAGGAGAGAG 1373
384 alThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400
1374 TAACATACTTGGGGTACAGTTTGGCGGACGGGACGGATGGCTGACGGAG 1423
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG 417
1424 GCACGGAAGAAACTGTAGTCCAGATACCGGCCCCACACACAGCCAAACA 1473
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
1474 AGTGAGAGAGTTTGGGGACAGCTGGATTTTGCAGACTGTGGATCCCGG 1523
434 LyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysLysLysGly 450
1524 GGTTCGCGACCTTAGCAGCCCCACCTCTACCCACTAACCAAGAAAGAGG 1573
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467
1574 GAATTCCTCTGGCTCTGAGCACCAGAGGCAATTTGATGCTATCAAAA 1623
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484
1624 GGCCCTCTGAGCGCACCTGCTGCGCCCTCCTGACGTAACATAAACCT 1673
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500
1674 TTACCCCTTATGTGATGAGCGTAGGGAGTAGCCCCGGGGAGTTTAAACC 1723
501 GlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLe 517
1724 CAAACTCTAGGACCATGGAGAGACCTGTGCTACCTGTCAAAAGACT 1773
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaVal 534
1774 CGATCCTGTAGCCAGTGGTGGCGGTATGCTGAAGGCTATCGCAGCTG 1823
534 aAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550
1824 TGCCCACTGTGTAAGGACCGTGAATAATGACTTTGGGACAGAATAATA 1873
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProAs 567
1874 ACTGTAATAGCCCCCATCGTGGAGACATCGTTCGCGACCCCCAGA 1923
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeu 584
1924 CGATGGATGACCAACGCCCGCATGACCCACTATCAAAAGCCTGCTTCTCA 1973
584 hrCluArgValThrPheAlaProAlaLeuAsnProAlaThrLeu 600
|||||

1974 CAGAGAGGGTCAGTTGCTGCCACAGCCGCTCTCAACCTGCCACTCTT 2023
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617
2024 CTGCTGAAGAGACTGATGAACAGTACTCATGATGTCATCAACTATT 2073
617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634
2074 GATTGAGGAGACTGGGTCGCAAGACCTTACAGACATACCGCTGACTG 2123
634 LyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650
2124 GAGAAGTGTAACTGTTGCTCAGTACGGAAGCAGCTATGTGGTGAAGGT 2173
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667
2174 AAGAGGATGCTGGGCGGCGGTGGTGGAGCGGACCCGACGATCTGGGC 2223
667 aserSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684
2224 CAGCAGCTCGCGGAAGAACTTACGCACAAAGGCTGAGCTCATGGCCC 2273
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700
2274 TCAGCAGGCTTTCGCGCTGGCGGGAAGAAATCCATAACATTTATACA 2323
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717
2324 GACAGCAGTATGCTTTGGGACTGCACAGTACACGGGGCCCATCTATAA 2373
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734
2374 GCAAGGGGGTGGCTTACCTCAGCAGGAGGGAATAAAGAACAAAGAGG 2423
734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750
2424 AAATCTTAAGCTATTAGAAGCCTTACATTTGCCAAAAGGCTAGCTATT 2473
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767
2474 ATACACTGTCTGGACATCAGAAAGCCAAAGATCCCATATCCAGAGGAA 2523
767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784
2524 CCAGATGGCTGACCGGTTGCCAAGCAGGAGCCCGGAGGTGTTAACCTTC 2573
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800
2574 TGCCTATGATAGAAACACCCAAAGCCCAAGACCCGACGACAGTACACC 2623
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817
2624 CTAGAAGACTGGCAAGAGATAAAAGATAGACCACTTCTGAGACTCC 2673
817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834
2674 GGAAGGACCTGCTTATACCTCAGATGGGAAGAAATCCTGCCCCACAAG 2723
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850
2724 AAGGTTAGATATGTCCAACAGATACATCGTCTAACCCACCTAGGAAC 2773
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867
2774 AAACACCTGCAGCAGTGGTGGTGAAGAACTCTCTATCATGTTCTGAGGCT 2823
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884
2824 ACCAGGAGTGGCTGATTGGTGGTGAACACTGTGTGCTGCGCTGCCAGCTGG 2873
884 alAsnAlaAsnProSerArgIleProGlyLysArgLeuLeuArgGlySer 900
2874 TTAATGCTAATCCTTCCAGATACTCCAGAAAGAGACTAAGGGGAAGC 2923

901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917
|||||
2924 CACCACGGCGTCACTGGGAAGTGGACTTCACCTAGGTAAAGCGGCTAA 2973
917 sTyGlyAsnLysTrpLeuLeuValPheValAspThrPheSerGlyTrpV 934
|||||
2974 ATACGGAACAAATATATCTATTTGGTTTGTAGACACCTTTTCAGATGG 3023
934 alGluAlaTyProThrLysLysGluThrSerThrValValAlaLysLys 950
|||||
3024 TAGAGGCTTATCTTAAAGAGACTTCAACCGTGGTGGCTAAGAA 3073
951 IleLeuGluLysPheProArgPheGlyIleProLysValIleGlyLe 967
|||||
3074 ATACTGGAGAAATTTTCCAGATTGGGAATACCTAAGTATAGGTC 3123
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984
|||||
3124 AGACAATGGTCCAGCTTTCGTTGCCAGGTAAAGTCAGGAGCTGGCCA 3173
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyArgProGlnSerSer 1000
|||||
3174 TATTTGGGATGATTGGAACACTGCATTGTGCATACAGACCCCAAGCTCA 3223
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017
|||||
3224 GGACAGGTAGAGAGGATGAATAGAACCATTAAGAGACCTTACTAAT 3273
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheVal 1034
|||||
3274 GACCGGGAGCTGGCGTTAATGATTGGATAGCTCTCTGCGCTTTGTGC 3323
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrglu 1050
|||||
3324 TTTTGGGTAGGAACACCCCTGGACAGTTGGGCTGACCCCTATGAA 3373
1051 LeuLeuTyrglyGlyProProProLeuAlaGluIleAlaPheAlaHis 1067
|||||
3374 TTACTCTACGGGGACCCCCCATTTGGTAGAAATGCTTCCTCGTACATAG 3423
1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084
|||||
3424 TGTGACGCTGCTCTTCCAGCCCTTTGTTCTCTAGGCTCAAGGCACTTG 3473
1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrs 1099
|||||
3474 AGTGGGTGAGACACAGCGTGGAGCAACTCCGGAGGCGCTACTCAGGA 3523
1100 GlyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTy 1116
|||||
3524 GGAGGAGCTTGCAGATCCACATCGTTTCCAAGTGGGAGATTCAGTCTA 3573
1116 rValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyPro 1133
|||||
3574 CGTTAGACGCCCGTGCAGGAACCTTCAGACTCGGTGGAAGGCGCCCTT 3623
1133 yrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIlePro 1149
|||||
3624 ATCAGCTACTTTTGACCACACACCGGCTGCTGAAGTGAAGAAATCTCC 3673
1150 ThrTrpIleHisAlaSerHisValLysProAlaProProAspSerG 1166
|||||
3674 ACCTGGATCCATGCATCCACCTTAAGCGCGGCCACCTCCCGATTCGGG 3723
1166 yTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgL 1183
|||||
3724 GTGGAAGCGGAAAGACTGAAATCCCTTAAAGCTTCGCTCCATCCGCG 3773
1183 euValProTyrsSerAsnAsnSer 1191
:::|||||
3774 TGGTTCCTTACTCTGCAATACTCC 3799

seq_name: gb_pat:AX052634
seq_documentation_block: 7362 bp DNA PAT 12-JAN-2001
LOCUS AX052634
DEFINITION Sequence 19 from Patent WO0071726.
ACCESSION AX052634
VERSION AX052634.1 GI:12226824
KEYWORDS
SOURCE
ORGANISM
porcine endogenous retrovirus.
porcine endogenous retrovirus
Viruses; Retroviral viruses; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
REFERENCE
1 (bases 1 to 7362)
AUTHORS
Federspiel, M.J.
TITLE
Methods to inhibit infectious agent transmission during
xenotransplantation
JOURNAL
Patent: WO 0071726-A 19 30-NOV-2000;
Mayo Medical Ventures (US)
FEATURES
Location/Qualifiers
1..7362
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
BASE COUNT 1997 a 1821 c 1881 g 1663 t
ORIGIN
alignment_scores:
Quality: 6095.50 Length: 1192
Ratio: 5.174 Gaps: 1
Percent Similarity: 98.826 Percent Identity: 96.980
alignment_block:
US-09-171-553B-5 x AX052634
Align seg 1/1 to: AX052634 from: 1 to: 7362
1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17
|||||
2823 GGAAGACGGGGTTCGGACCCCTCCCGAGCCGAGGTTAACTTTGAAGGT 2872
17 IGLuGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34
|||||
2873 GGAGGGCAACCACTGAGTTCTCTGGTTGATACCGGAGCGAAACATTGAG 2922
34 alLeuLeuGlnProLeuGlyLysLeuLysGluLysLysSerTrpValMet 50
|||||
2923 TGCTACTACAGCCATTAGGAAACTAAAGATAAAAAATCCTGGGTGATG 2972
51 GlyAlaThrGlyGlnArgGlnTyTrpThrArgArgThrValAs 67
|||||
2973 GGTGCCACAGGCAACACAGTATCCATGGACTACCCGAGAGAACAGTTGA 3022
67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysp 84
|||||
3023 CTTGGAGTGGGAGGGTAAACCCACTCGTTTCTGTCTCATACCTGAGTGCC 3072
84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100
|||||
3073 CAGCACCCCTCTTAGGTAGAGACTTATTGACCAAGATGGGAGACAAATT 3122
101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117
|||||
3123 TCTTTTGAACAAGGAAACACAGAGTGTCTGCAATTAACAACCTATCAC 3172
117 rValLeuThrLeuGlnLeuAspAspGluTyArgLeuTySerProGlnV 134
|||||
3173 TGTCTTGACCTCCCAATTAGATGACGAATATCGACTATACTCTCCCTAG 3222
134 alLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150
|||||
3223 TAAAGCCTGATCAAAATATACAAATCTGGTTGGAAACAGTTTCCCAAGCC 3272
151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProGlnVa 167

||||| 3273 TGGCAGAAACCGCAGGATGGTTGGCAAGCAAGTTCCTCCACCAAGT 3322
167 lileGlnLeuLysAlaSerAlaThrProValSerValArgGlnThrProL 184
3323 TATTCACCTGAAGGCGATGCGCACACAGTGTCACTCAGCAGTACCCCT 3372
184 euserArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200
3373 TGAGTAAGAGAGCTCAAGAGGATTCGGCCGCATGTCCAAAGATTATC 3422
201 GlnGlnGlyIleLeuValProValGlnSerProThrPasnThrProLeuLe 217
3423 CAACAGGCATCCTAGTTCCTGTCCTCAATCTCCCTGGATATCTCCCTGCT 3472
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234
3473 ACCGGTTAAGAGCTGGGAGCTAATGACTATCGACCATGACGACACTTGA 3522
234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250
3523 GAGAGTCAATAAAGCGGTGAGGATATACACCAACAGTCCCGAACCT 3572
251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValLe 267
3573 TATAACCTCTTGTGTGCTCTCCACCCCAACGAGCTGGTATACAGTATT 3622
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284
3623 GGACTTAAGAGGATGCTTTTCTGCCTGAGATTACACCCCACTAGCCCAAC 3672
284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
3673 CACTTTTTCCTTCCGATGGAGAGATCCAGGTACGGGAAGAACCGGCGAG 3722
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePhe 317
3723 CTCACCTGGACCCGCTGCCCCAAGGGTTCAAGAACTCCCGCAGCATCTT 3772
317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
3773 TGACGAGAGCCCTPACACAGAGACTCGCCCACTCAGGATCAACACCCCTC 3822
334 InValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350
3823 AGGTGACCTCTCTCCAGTACGTGGATGACTCTCTTGGCGGAGCCACC 3872
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAs 367
3873 AAACAGGACTGCTTGAAGGCACGAGGCACTTCTGGAATTGCTGA 3922
367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
3923 CTTAGCTACAGAGCTCTGCTTGAAGCCCGCAGATTTCAGGAGAGAGG 3972
384 alThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400
3973 TAACATACCTGGGTACAGTTTGGCGGACGCGGCGGATGGCTGACGGAG 4022
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG 417
4023 GCACGGAAGAAACTGTAGTCCAGATACGGCCCAACACAGCCAAACA 4072
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
4073 AGTGAGAGAGTTTGGGACAGCTGGATTTCAGAGCTGTGATCCCGG 4122
434 lypPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGlnLysGly 450
4123 GGTTCGACCTTAGCAGCCCACTCTACCCACTAACCAAGAAAGAGG 4172
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysL 467
|||||

4173 GAATTCCTCTGGCTCTCTGAGCACCAAGGCATTTGATGCTATCAAAA 4222
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProP 484
4223 GGCCTCTGCTGAGCGCACCTGCTCTGGCCCTCCCTGACGTAACCTAACCCCT 4272
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500
4273 TTACCTTTTATGTGTGATGAGCGTAAGGAGTAGCCCGGGAGTTTAAACC 4322
501 GlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLe 517
4323 CAACCTCTAGGACCAATGGAGAGACCTGTTCCTACTCTCAAGAGAGCT 4372
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaAlav 534
4373 CGATCTGTAGCCAGTGGTTGGCCCGTATGCCCTGAAGGCTATCGCAGCTG 4422
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550
4423 TGCCCATACTGGTCAAGGAGCTGACAAATTTGCTTTGGGACAGAAATA 4472
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567
4473 ACTGTATAGCCCTCCATGCGTTGGAGACATCGTTCCGCGAGCCCCAGA 4522
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeu 584
4523 CCGATGATGACCAACGCCCGCATGACCCACTATCAAAAGCTCTCTCTCA 4572
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600
4573 CAGAGAGGTCAGTTCGCTCCACCGCCCTCTCAACCTGCCACTCTT 4622
601 LeuProGluGlnThrAspGluProValThrHisAspCysHisGlnLeuLe 617
4623 CTGCTGACAGACTGATGAACCACTGACTCATGATGCCATCACTATT 4672
617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634
4673 GATTGAGGAGACTGGGTCGCGCAAGGACCTTACAGACATACCGCTGACTG 4722
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650
4723 GAGAAGTGTAACTTGGTTCACGAGGAGCAGCTATGTTGGTGAAGGT 4772
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667
4773 AAGAGGATGCTGGGCGCGCGTGGTGACGGGACCGCCGACGATCTGGGC 4822
667 aSerSerProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684
4823 CAGCAGCTGCGGAGAGAACTTTCAGACAAAAGGCTGAGCTCATGGCC 4872
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700
4873 TCAGCAAGCTTTTCGGCTGCGGCAAGGAAATCCATAAACATTATACA 4922
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717
4923 GACAGCAGGTATGCTTTGCGACTGACACGCTACACGGGGCCATCTATAA 4972
717 sGlnArgGlyLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734
4973 GCAAGGGGTTGCTTACCTCAGCAGGAGGGAATAAAGAACAAAGAGG 5022
734 luIleLeuSerLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750
5023 AAATTCATAGCCTATTAGACCTTACATTTGCCAAAAGGCTAGCTATT 5072
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767
5073 ATACACTGCTCTGACATCATGAAAGCCAAAGATCCCATATCCAGAGGAA 5122
|||||

```

767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784
|||||
5123 CCAGATGGCTGACCGGGTTGCCAGCAGGAGCCCGAGGGTTAAACCTTC 5172
|||||
784 euProfilelleGluThrProLysAlaProGluProArgArgGlnTyrThr 800
|||||
5173 TGCCTATGATAGAAACACCCAAAGCCCGAGAACCCCGGACGACAGTACACC 5222
|||||
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817
|||||
5223 CTAGAACACTGGCAAGAGATAAAAGATAGACACAGTCTCTGAGACTCC 5272
|||||
817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834
|||||
5273 GGAAGGACCTGCTATACCTCAGATGGGAAGGAATCTCTGCCACCAAG 5322
|||||
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850
|||||
5323 AAGGGTTAGAAATATGTCCAACAGATACATCGTCTAACCCACCTAGGAAC 5372
|||||
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867
|||||
5373 AAACACCTGCAGCAGTGGTCAGAACATCTCTTATCATGTCTTGAGGCT 5422
|||||
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884
|||||
5423 ACCAGGAGTGGCTGATTCGGTGGTCAAACTGTGTGCCCTGCCAGCTGG 5472
|||||
884 alAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900
|||||
5473 TTAATGCTAATCTCTCCAGAAATACCTCCAGCAAGAGACTAAGGGGAAGC 5522
|||||
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917
|||||
5523 CACCCAGGCGCTCACTGGGAAGTGGACTTCACCTCAGGTAAGCCGGCTAA 5572
|||||
917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934
|||||
5573 ATACGGAACAATAATCTATTTGTTGTAGACACCTTTTCAGGATGGG 5622
|||||
934 alGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950
|||||
5623 TAGAGGCTTATCTCTACTAAGAAAGAGACTTCAACCGTGGTGTAAAAA 5672
|||||
951 IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967
|||||
5673 ATACTGGAGGAATTTTCCAAGATTGGGAATACCTAAGGTAATAGGGTC 5722
|||||
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984
|||||
5723 AGACAATGGTCCAGCTTCCTGCCAGGTAACTCAGGACCTGGCCCAAG 5772
|||||
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000
|||||
5773 TATTGGGGATTGATTGGAACTGCATTGTGCATACAGACCCCAAGCTCA 5822
|||||
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017
|||||
5823 GGACAGGTAGAGAGGATGAATAGAACCATTTAAAGAGACCTTACTAAAT 5872
|||||
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValI 1034
|||||
5873 GACCGGGAGACTGGCGTTAATGATTTGGATAGCTCTCTGCCCTTTGTAC 5922
|||||
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050
|||||
5923 TTTTGGGGTTAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATGAA 5972
|||||
1051 LeuLeuTyrGlyGlyProProProLeuAlaGluIleAlaPheAlaHisSe 1067
|||||
5973 TTACTACGGGGGACCCCGCCCATTTGGTAGAAATTCCTCGGTACATAG 6022

```

```

1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084
|||||
6023 TGTGACGTGCTGCTTTCCAGCCTTTGTCTTAGGCTCAAGGCACCTTG 6072
|||||
1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSer... 1099
|||||
6073 AGTGGGTGAGACACACGAGCGTGGAGGCAACTCCGGGAGCGCTACTCAGGA 6122
|||||
1100 GlyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTy 1116
|||||
6123 GGAGGAGACTTGCAGATCCACATCGTTTCCAAAGTGGGAGATTCAGTCTA 6172
|||||
1116 rValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProT 1133
|||||
6173 CGTTAGACGCCACCGTGCAGGAACCTCGAGACTCGGTGGAAGGCCCTT 6222
|||||
1133 yrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIlePro 1149
|||||
6223 ATCAGCTACTTTTGACCACCAACGGCTGTGAAAGTCGAAGGAATCTCC 6272
|||||
1150 ThrTrpIleHisAlaSerHisValLysProAlaProProAspSerGl 1166
|||||
6273 ACCTGGATCCATGCATCCACGTTAAGCGCGGCCACCTCCCGATTCCGG 6322
|||||
1166 yTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgL 1183
|||||
6323 GTGGAAAGCCGAAAGACTGAAATATCCCTTAAGCTTCGCTCCATCCGCG 6372
|||||
1183 euValProTyrSerAsnAsnSer 1191
|||||
6373 TGGTCTCTTACTCTGTCAATAACTCC 6398
|||||
seq_name: gb_om:AF038600
seq_documentation_block:
LOCUS AF038600 8132 bp mRNA MAM 21-MAY-1998
DEFINITION Sus scrofa porcine endogenous retrovirus PERV-MSL, complete
sequence.
ACCESSION AF038600
VERSION AF038600.1 GI:3133301
KEYWORDS
SOURCE
ORGANISM Sus scrofa
pig.
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
AUTHORS Akiyoshi,D.E., Denaro,M., Zhu,H., Greenstein,J.L., Banerjee,P. and
Fishman,J.A.
TITLE Identification of a full-length cDNA for an endogenous retrovirus
of miniature swine
JOURNAL J. Virol. 72 (5), 4503-4507 (1998)
MEDLINE 98216827
REFERENCE
2 (bases 1 to 8132)
AUTHORS Fishman,J.A.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1997) Infectious Disease Unit, Massachusetts
General Hospital, 55 Fruit Street, Boston, MA 02114, USA
REFERENCE
3 (bases 1 to 8132)
AUTHORS Fishman,J.A.
TITLE Direct Submission
JOURNAL Submitted (15-MAY-1998) Infectious Disease Unit, Massachusetts
General Hospital, 55 Fruit Street, Boston, MA 02114, USA
REMARK Sequence update by submitter
COMMENT On May 16, 1998 this sequence version replaced gi:3116441.
FEATURES
source
1..8132
/organism="Sus scrofa"
/strain="miniature swine"
/db_xref="taxon:9823"
/cell_type="lymphocytes"
/notes="porcine endogenous retrovirus PERV-MSL"
1..564
585..2159
LTR
CDS

```

```

/codon_start=1
/product="gag protein"
/protein_id="AAC16763.1"
/db_xref="GI:3116442"
/translation="MGQVTTPSLTLDHWTVEKSPAHNLSVQVKKGPMQWTCFVSEMP
TFDVGPSGTFNSEILLAVKAVIFQTCGSHDDQEPYILLTWODLAEDPPVWVKPWLN
KPKPKPRILALGKNAEKVKSPHYIELEPPAMPPEOSPPPPYLAOGAARG
FPAPGAAVEGAAGTASRGATPERTDEIATPLRTYPTPGQLOPLQWFFS
ADLYNNKTHNPESDQRLTGLVSLMFSDQMTDCCOQLQTLFTFERERILLER
KRVNCGADGRTPLONEIDMGPLTRPGWDNTAEGRESLKIVROALVAGLRASRRP
TNLAKRYEVMQGNPEPSPVLELLAEFRVYTFPOPTSEAKASVALAFIGOSADIR
KKQLEGLQEAERDLVKAEEKYIKRETEERDOKREEREEREERNKRNKQENLT
KILAAVVEGSEKTEREDFRKINSRQSGNLGNRTPLDKQOCAYCKEKGHWARNCPK
KGNKGPRILALEEDKD"
2307.5744

```

CDS

```

/codon_start=1
/product="pol protein"
/protein_id="AAC16764.1"
/db_xref="GI:3133302"
/translation="MGATGQQQYPTWTRRTVDLGVGRVTHSLVIPCAPLLGRDLL
TKMAQISFEOGKEPVSANNKPTVLTLQDDDEYRLYSPLVKPDONIQWLQFQPOAW
AFTAGAGLAKOVPOVIOLKASATPVSVQYPLSKEAQEGIRPHVORLIQOQGIKVPVO
SPWNTPLLVRKPGTNDYRPVQDLREVNKRQVDIHTVVPNPNLICALPPQRSWTVL
DLKDAFFKLRLHPTSQLFAFEMRDEGTGOLTWRLPQGGKNSPTFFDGLARHDL
ANFRHQPTVLLQYVDDLLLAGATKQDCLEGTALLLESLDLGYRASAKKAOICRE
VYLYGSLRDGQWLTARAKTVQIAPPTAKQKREFLGTAGFCRLWTPGFGATLAAP
LPLTKREGEFSWPEHQAFDAIKKALLSAPALPDVTKPTFLYVDRKGVARGVL
TQTLGPRRPVAVLSKLLDPVAGSWPICLKATAAVAILVKDADKLTGONITVIAPHA
LENIVRQPDWRMTNARMTHYQSLLLTERTVFPAPPAALPATLLPEETDEPVTDOCHO
LLEETGPRKDLTDIPLTGEVLTFTDGSYVVEGKRMAGAAVVDGTRTIWASSLPGG
TSAQKELMALQALRLAEGKSNINYTDRIYAFATAHVGAIYKQGLLTSAGREIKN
KEBILLSLEALHPLKRLAIHCPGHOKAKDLISRGNMADRAVKAQAAQGNLLPIIEM
KPAPEPRVOTLEDQWIEIKKIDQFSETPEGTCTYSDGKELIPKLEGLVYQIHLTH
LGTKLQQLVRTSPYHVLRLPGVADSVVKHCVPCLVNPANPSRMPGKRLGRSHPGAH
WEYDFTEVPKAKGKXLLVFDVTFESGVEAYPTKETSTVVAKKILEIFRFGIPK
VIGSDNGPAFVAVHVSQGLAKILGIDMKLHCAYRQSSGQVERMNTIKETLLKLAET
GNDMIALPLFVLVRNTPGQPLTYELLYGGPPPLVEIASVHSADVLISQPLFSR
LKALEWRQRMQRLREAYSGGDDQIPIHRFOVQDSVYVRRHAGNLETRWKGPIYLV
LTPYAVKVEGISTWIHASHVRPAPPDPSGWAKEENTENPLKRLHRVVPYSVYNLSD"
5620.7536

```

CDS

```

/codon_start=1
/product="env protein"
/protein_id="AAC16765.1"
/db_xref="GI:3133303"
/translation="MHPTLNRRHLPIRGKPKRLKIPLSFASITAMFLTISITQTNM
RIGDSLNSHKLTLTIDTSGTGINNTQGEAPLGTWMPDLYVCLRSVIPSLTSP
DILHANGFYVCPGPNNGKHCNPDRDFCKQWNCVTSDNGYKWPSTQODRVSFYVG
TYTSQGFNLTWIRTSKSPSDDLKLSFTKQKQENTILKWNGSGWGVYVG
SGKQPSILIRUKINOLEPMAIGNTVLTGORPTQPGFSSNTSSSDTESST
TKWAKLSLIQAFQALNSTTPEATSSCWLCLASGPPYYEGMARKEFNVTKEHRDQ
CTWGSQNKLTLEVSGKTCIGKVPSSHQHCNHTFAFNQTSOYLVPYDWRWACN
TGLTPCVSTLVNOTKDFCIMOIVPRVYYPEKALIDEYDYNHRQKREPTSLTIAV
MLGLVAAVGTCGTAALVTCPOLETGLSNLHRIVTEDLQALEKSVNSLESLTISE
VVLONRRGLDLLFLKGGCLVALKECCFYVDHSGAIRDSMNKLRLEKRRERETT
QGWFEWGNLSMLATLLSALTGLVILLVLLLTIVGPCIINKLFIPIREKISAVQIMVL
ROOYQSPSSREAGR"
7518.7976

```

```

LTR          2247 a 1977 c 2037 g 1871 t
BASE COUNT
ORIGIN

```

```

alignment_scores:
  Quality: 6087.50      Length: 1191
  Ratio: 5.168          Gaps: 1
  Percent Similarity: 98.908  Percent Identity: 96.725

```

```

alignment_block:
US-09-171-553b-5 x AF038600

```

```

Align seg 1/1 to: AF038600 from: 1 to: 8132

```

```

1 GlyArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17
|||||

```

```

2160 GGGAGAGCGGGTTCGGACCCCTCCCGAGCCCGAGGTAACCTTTGAAGGT 2209
17 lGluGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34
|||||
2210 GAGGGGCAACCAAGTTGAGTTCTCTGTTGATACCGGAGCGAAACATTCAG 2259
34 allLeuLeuGlnProLeuGlyLysLeuLysGluLysLysSerTrpValMet 50
|||||
2260 TGTACTACAGCCATTAGGAAAACTAAAAAGATAAAAAATCCTGGGTGATG 2309
51 GlyAlaThrGlyGlnArgGlnTrpProTrpThrArgArgThrValAs 67
|||||
2310 GGTGCCACAGGGCAACACAGTAGTATCCATGAGCACCAGAGAACAGTTGA 2359
67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysP 84
|||||
2360 CTTGGGAGTGGGACGGGTAAACCCACCTCGTTTCTGGTCATACCTGAGTGCC 2409
84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100
|||||
2410 CAGCACCCCTCTTAGGTAGAGACTTATTGACCAAGATGGGAGCACAAAT 2459
101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117
|||||
2460 TCTTTTGAACAGGGAAACAGGAAGTGTCTGCAAAATAACAACTATCAC 2509
117 rValLeuThrLeuGlnLeuAspGluValArgLeuValSerProGlnV 134
|||||
2510 TGTGTTGACCTCCAAATTAGATGACGAATATGACTATATCTCTCCCTAG 2559
134 allLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150
|||||
2560 TAAAGCCTGATCAAAATATACAAATCTCTGTTGGAACAGTTTCCCCAGGC 2609
151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnVa 167
|||||
2610 TGGCAGAACCCGAGGATGGGTTTGGCAACCAAGTTCCCCCAAGT 2659
167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTrpProL 184
|||||
2660 TATTCAACTGAAGGCCAGTGCCACAGCAGTGTCTCAGTCAGACAGTACCCCT 2709
184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200
|||||
2710 TGAGTAAAGAGCTCAAGAGGAATTCGGCCGCTGTCACAAAGATTAAATC 2759
201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217
|||||
2760 CAACAGGCATCTAGTCTCTGTCCTCAATCTCCCTGGAATACCTCCCTGCT 2809
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234
|||||
2810 ACCGGTTAGAAAGCCTGGGACTAATGACTATCGACCAAGTACAGGACTTGA 2859
234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250
|||||
2860 GAGAGGTCATAAACGGGTGAGGATATACCCCAACAGTCCCGAACCCCT 2909
251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValLe 267
|||||
2910 TATAACCTCTTGTGCTCTCCACCCACGAGGAGCTGGGTATACAGTATT 2959
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284
|||||
2960 GGACTTAAAGGATGCCCTCTCTCTGCTGAGATTACACCCCACTAGGCAAC 3009
284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
|||||
3010 CACTTTTGGCTTCGAATGAGAGATCCAGGTACGGGAAGAACCCGGGACG 3059
301 LeuThrTrpThrArgLeuProGlnGlnGlyPheLysAsnSerProThrIlePh 317
|||||
3060 CTCACCTGGACCGACTGCCCCAGGGGTTCAAGAACTCCCCGACCATCTT 3109

```

317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
|||||
3110 TGACGAAGCCCTACACAGAGACCTGGCCAACTTCAGGATCCACACCCCTC 3159
|||||
334 InValThrLeuLeuGlnTyrValIAspAspLeuLeuAlaGlyAlaThr 350
|||||
3160 AGGTGACCCCTCTCCAGTAGTGATGACCTGCTTCTGGGGAGGCCACC 3209
|||||
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuGluLeuSerAs 367
|||||
3210 AAACAGACTGCTTAGAAGGACCAAGGCACACTACTGCTGGAATTTGCTGA 3259
|||||
367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
|||||
3260 CCTAGGCTACAGAGCCCTCTGCTAAGAAGGCCAGATTTCAGGAGAGAGG 3309
|||||
384 alThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400
|||||
3310 TAACATACCTTGGGGTACAGTTTACGGGACGGGACGATGGCTCACGGAG 3359
|||||
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG1 417
|||||
3360 GCACGGAGAAACTGTAGTCCAGATACCGGCCCCCAACCCACGCCAAACA 3409
|||||
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
|||||
3410 AATGAGAGAGTTTTTGGGGACAGCTGATTTGCAGACTGTGGATCCCGG 3459
|||||
434 lyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450
|||||
3460 GGTTCGGACCTTAGCAGCCCACTCTACCCGCTAACCAAGAAAGGG 3509
|||||
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467
|||||
3510 GAATTCCTCGGGCTCTGAGCACCAGAAGGCATTTGATGCTATCAAAA 3559
|||||
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484
|||||
3560 GGCCCTCTGAGCGCACCTCTCTGGCCCTCCCTGACGTAACCTAACCCCT 3609
|||||
484 hethrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500
|||||
3610 TTACCCCTTTATGTGGATGACCGTAAGGGAGTAGCCCGGGGAGTTTTAAC 3659
|||||
501 GlnThrLeuGlyProTrpArgArgProValAlaIaIaTyrIleuSerLysLe 517
|||||
3660 CAACCCCTAGACCATGGAGAAGACCTGTGCGCTACCTGTCAAGAAAGCT 3709
|||||
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaAlay 534
|||||
3710 CGATCCTGTAGCCAGTGGTTGGCCCATATGCTGAAGGCTATCGCAGCTG 3759
|||||
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550
|||||
3760 TGGCCATCTACTGGTCAAGGACGCTGACAAATTGACTTTGGACAGAAATA 3809
|||||
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567
|||||
3810 ACTGTAATAGCCCCCATGATTTGGAGAACATCGTTTCGGCAGCCCCCAGA 3859
|||||
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuT 584
|||||
3860 CCGATGGATGACCAACGCCCGCATGACCCACTATCAAGCCCTCTCTCA 3909
|||||
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600
|||||
3910 CAGAGAGGGTCAAGTTCGCTCCACCACGCGCTCTCAACCCCTGCCACTCT 3959
|||||
601 LeuProGluThrThrAspGluProValThrHisAspCysHisGlnLeuLe 617
|||||
3960 CTGCCCTGAAGAGACTGATGAACCAAGTGACTCATGATTGCCATCAACTATT 4009

617 uileGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634
|||||
4010 GATTGAGGAGACTGGGGTCCGCAAGACCTTACAGACATACCGCTGACTG 4059
|||||
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650
|||||
4060 GAGAAGTGCTAACTGTTCACTGACGGAGCAGCTATGTGTGGTGAAGGT 4109
|||||
651 LysArgMetAlaGlyAlaValValAspGlyThrArgThrIleTrpAl 667
|||||
4110 AAGAGGATGGCTGGGGCGGGTGGTGGACGGGACCCGCCACGATCTGGGC 4159
|||||
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAla 684
|||||
4160 CAGCACCTTGGCGGAGGAACTTCAGCACAAAAGGCTGAGCTCATGGCCC 4209
|||||
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700
|||||
4210 TCACGCAAGCTTTGCGCTGGCGGAGGGAATCCATAAACATTTATACG 4259
|||||
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717
|||||
4260 GACAGCAGGTATGCCCTTTCGACTGCACACGTACATGGGGCCATCTATAA 4309
|||||
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734
|||||
4310 ACAAAAGGGGTTGCTTACCTCAGCAGGAGGGAATAAAGAACAAAGAGG 4359
|||||
734 luileLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750
|||||
4360 AAATTTCAAGCCTATTAGAAGCGGTACATTTACCAAAAAGGCTAGCTATT 4409
|||||
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767
|||||
4410 ATACACTGCTCGGACATCAGAAAGCTAAAGATCTCATATCCAGAGGAAA 4459
|||||
767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784
|||||
4460 CCAGATGGCTGACCGGGTTGCCAAGCAGCAGCCCGAGGGTGTAACTTC 4509
|||||
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800
|||||
4510 TGCCCTATAATAGAAATGCCAAAGCCCGAGAACCCAGACGACAGTACACC 4559
|||||
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817
|||||
4560 CTAGAAGACTGGCAAGAGATAAAAGATAGACCAGTCTCTCTGAGACTCC 4609
|||||
817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834
|||||
4610 GGAAGGACCTGCTATACCTCAGATGGGAAGGAATCTGCCCCACAAG 4659
|||||
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850
|||||
4660 AAGGGTTAGATATGTCCAACAGATACATCGTCTAACCCACCTAGGAAC 4709
|||||
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867
|||||
4710 AAACACCTGCGACGAGTGTGTCAAGACATCCCCCTTATCATGTTCTGAGGCT 4759
|||||
867 uProGlyValAlaAspSerValLysHisCysValProCysGlnLeuV 884
|||||
4760 ACCAGAGTGGCTGACTCGGTGGTCAACATTTGTGTGCCCTGCCAGCTGG 4809
|||||
884 alAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900
|||||
4810 TTAATGCTAACTCTTCCAGAAATGCTCCAGGGAAGAGACTAAGGGGAAGC 4859
|||||
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917
|||||
4860 CACCAGGCGCTCACTGGGAAGTGGACTTCACCTGAGGTAAAGCCGCTAA 4909
|||||
917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934

```
|||||
4910 ATACGGAACAACAAATACCTATTGGTTTGTAGACACCTTTTCAGATGGG 4959
934 aGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950
|||||
4960 TAGAGGCTATCTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAAGAAA 5009
951 ileLeuGluGluIlePheProArgPheGlyIleProLysValIleGlyse 967
5010 ATACTGGAGAAATTTTCCAAAGATTGGAAATACCTAAGGTAATAGGGTC 5059
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984
|||||
5060 AGACAATGTCAGCTTTGTTGCCACGTAAGTCAGGAGCTGCCAAGA 5109
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000
5110 TATTGGGATGATTGGAAACTGCATTGTGCATACAGACCCCAAGCTCA 5159
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017
5160 GGACAGGTAGAGAGATCAATAGAACCATTAAGAGACCCCTTACTAAAT 5209
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034
5210 GACCCGCGAGAGCTGGCTTAATGATTGGATAGCTCTCCTGCCCTTTGTC 5259
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050
5260 TTTTATAGGTTAGGAACACCCCTGGACAGATTGGGCTGACCCCTATGAA 5309
1051 LeuLeuTyrGlyGlyProProProLeuAlaGluIleAlaPheAlaHisSe 1067
5310 TTACTCTACGGGGACCCGCCCATTTGTTAGAAATGCTTCTGTACATAG 5359
1067 rIlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084
5360 TGCTGATGCTGCTTTTCCAGCCTTTGTTCTTAGGCTCAAGGCACTTG 5409
1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSer... 1099
5410 AGTGGGTGAGACAACGCGTGGAGGCACTCCGGAGGCGCTACTCAGGA 5459
1100 GlyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTy 1116
5460 GGAGGAGACTTGCAGATCCCATCGTTTCCAAAGTGGGAGATTTCAGTCTA 5509
1116 rValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProT 1133
5510 CGTTAGACGCCACCGTGGAGAAACCTCGAGACTCGGTGGAGGCGCCCTT 5559
1133 yrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIlePro 1149
5560 ATCTCGTACTTTTGACCACACCAACCGCTGTGAAAGTCGAAGGAATCTCC 5609
1150 ThrTrpIleHisAlaSerHisValLysProAlaProProProAspSerGl 1166
5610 ACCTGGATCCATCATCCACGCTTAACCGGCGCCACCTCCCGATTCGGG 5659
1166 yTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgL 1183
5660 GTGGAAAGCCGAAAGACTGAAATCCCTTAAGCTTCGCTCCATCCGCG 5709
1183 euValProTyrSerAsnAsn 1190
:::|||||
5710 TGGTTCCTTACTCTGTCAATAAC 5732
```

seq_name: gb_om:AF038599

seq_documentation_block:

LOCUS AF038599

DEFINITION Sus scrofa porcine endogenous retrovirus Tsukuba-1 mRNA, complete

sequence.

21-MAY-1998

MAM

```
ACCESSION AF038599
VERSION AF038599.1
KEYWORDS GI:3133300
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 8132)
AUTHORS Akiyoshi,D.E., Denaro,M., Zhu,H., Greenstein,J.L., Banerjee,P. and
Fishman,J.A.
TITLE Identification of a full-length cDNA for an endogenous retrovirus
of miniature swine
JOURNAL J. Virol. 72 (5), 4503-4507 (1998)
MEDLINE 98216827
REFERENCE 2 (bases 1 to 8132)
AUTHORS Fishman,J.A.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1997) Infectious Disease Unit, Massachusetts
General Hospital, 55 Fruit Street, Boston, MA 02114, USA
REFERENCE 3 (bases 1 to 8132)
AUTHORS Fishman,J.A.
TITLE Direct Submission
JOURNAL Submitted (15-MAY-1998) Infectious Disease Unit, Massachusetts
General Hospital, 55 Fruit Street, Boston, MA 02114, USA
REMARK Sequence update by submitter
COMMENT On May 16, 1998 this sequence version replaced gi:3116437.
FEATURES
Location/Qualifiers
1..8132
/organism="Sus scrofa"
/db_xref="taxon:9823"
/cell_line="Shimozuma-1; malignant lymphoma derived cell
line"
/note="porcine endogenous retrovirus Tsukuba-1;
spontaneously produced virions of endogenous retrovirus
type C of swine were used to infect a swine kidney cell
line"
1..592
593..2167
/codon_start=1
/product="gag protein"
/protein_id="AAC16761.1"
/db_xref="GI:3116438"
/translation="MGQVTTPLSLTDHTEVKSRAHLSVOVKGPWOTFCVSEWP
TFDVGWSEGFNSEIILAVKAIIFQIGPGSHPOEPIYLWQDLAEDPPWVPWLN
KRPKPRILALGKRNKHSABKVPSPHIYPIEIEPPAWPEPQSPVPPPIYLAQGAAG
PFAPGAPAVEGSPAGTSRRGATPERTDEIATLPTRTGPTPGGQLQYWPFS
ADLYNWKTNHPFSEDQRLTGLVESLFSHQPTWDCQQLQTLFTFEERILLEA
RKNVPGADRPRLQNEIDMGFPILTRPGMDYNTAEGRESLKIYQALVAGLRGASRRP
TNLAKREVMOGPNPPSVFLERLLEAFRRYTPEDPTSEAKOASVALAFIQSALDIR
KKLQRLGLEQAEALDLVKEAKYIYKRETEEREQKREEREEREERNRKQENLT
KILAAVVEGKNTREDRFRKIRSGPRQSGNLGNRTPLDKDQCAFCYCKERGHWRNCPK
KGNKGPRILALEEDKD"
2315..5751
/note="pol protein"
/codon_start=1
/pseudo
5627..7543
/codon_start=1
/product="env protein"
/protein_id="AAC16762.1"
/db_xref="GI:3116439"
/translation="MHPTLSRRHLPIRGSKPKRLKIPLSFASIAWFLTLSTTSNGM
RIGDSINSHKPLSLTWLITDSTGNTINNTGEPALGTWNPDLVCLRSVIPSITSPP
DILHAHFYVCPGPNNGKCNPRDFCKNCVTSNDGYKWKPTSOODRVSFSYVN
TYTSSQFNLYLTWRTGPKSCPSDLDKLSTFKGKQENILKWNMGSMGWYVGG
SGKQPSILITRLKINQLEPNMTALQRPPTQRPQGPSSNITSGDPTESST
TKMGAKLSLIQGAFOALNSTPPEATSSCLCLALGPPY YEGMARRGFNVTKHRDQ
CTWGSQNKLTLEVSGKCTCIGKVPVSHQHLNHTENAFNSETSOYLVPGVDRWACN
TGLTFCVSTLVFNQTKDCIMVQIVPRVYIYPEKALIDEYDYNRHRQREPSITLAV
MIGLGAAGVGTTAAALTPQOGLSLNLRIVTEDQLAKESVLSLESLTSLSE
VVLQNRGLDLFLKEGGLCVALKEECCFYVDHSGATDLSMKNLRERLEKRRKETT
OCWFEFGWNRSPWLATLLSALTGPLIVLLLLLTGVCPIINKLIATIRIRISAVQIMVL
RQOQYSPSSREAGR"
```


LTR 7548..8132
 BASE COUNT 2247 a 1980 c 2030 g 1875 t
 ORIGIN

alignment_scores:
 Quality: 5979.50 Length: 1193
 Ratio: 5.076 Gaps: 3
 Percent Similarity: 98.743 Percent Identity: 96.479

alignment_block:

US-09-171-553B-5 x AF038599 ..

Align seg 1/1 to: AF038599 from: 1 to: 8132

```

1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17
|||||
2168 GGGAGACGGGGTTCGGACCCCTCCCGAGCCCGAGGGTAACCTTTGAAGGT 2217
17 lGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34
|||||
2218 GGAGGGCAACCAAGTTGAGTTCTCGTTGATACCGGAGGGAACATTCAG 2267
34 allLeuLeuGlnProLeuGlyLysLysLysGluLysLysSerTrpValMet 50
|||||
2268 TGTCTACTACAGCCATTAGGAATACTAAAGATAAAAAATCCTGGGTGATG 2317
51 GlyAlaThrGlyGlnArgGlnTyrProTyrThrArgGlnThrValAs 67
|||||
2318 GGTGC.ACAGGGCAACACAGTATCCATGGACTACCCGAAG.ACAGTTGA 2365
67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysp 84
|||||
2366 CTTGGGAGTGGGACGGTAAACCCACATCGTTCTCGGTATACCTGAGTGCC 2415
84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100
|||||
2416 CAGCACCCCTCTTAGGTAGAGACTTATTGACCAAGATGGGAGCACAAAT 2465
101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117
|||||
2466 TCCTTTGACACAGGGAACAGAGGTCTGCAAAATACAAACCTATCAC 2515
117 rValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnV 134
|||||
2516 TGTGTTGACCTCCCAATTAGATGACCAATATCGACTATACTCTCCCTAG 2565
134 allysProaspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150
|||||
2566 TAAAGCCTGATCAAAATATACAAATCTGGTTGSAACAGTTTCCCAAGCC 2615
151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnVa 167
|||||
2616 TGGGCAGAACCCGACGGATGGTTGGCAAGCAAGTTCCCCCAAGT 2665
167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184
|||||
2666 TATTCAACTGAAGCCAGTGCACACACAGTGTGAGTCAGTCAGACAGTACCCT 2715
184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeulle 200
|||||
2716 TGAGTAAAGAAAGCTCAAGAAAGAAATTCGGCCGCATGTCCAAAGATTAATC 2765
201 GlnGlnGlyIleLeuValProValGlnSerProTyrAsnThrProLeuLe 217
|||||
2766 CAACAGGGATCCTAGTCTCTGTCCAACTCCCTGGAATATCCTCCCTGCT 2815
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234
|||||
2816 ACCGGTTAGAAAGCCTGGGACTTAATGACTATCGACAGTACAGGACTTGA 2865
234 rgGluValAsnLysArgValGlnAspTleHisProThrValProAsnPro 250
|||||

```

```

2866 GAGAGGTCATAAAACGGGTGCAGGATATACACCAACAGTCCCGAACCT 2915
251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValle 267
|||||
2916 TATAACCTCTTGTGTCTCTCCACCCCAACGGAGCTGGTATACAGTATT 2965
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnp 284
|||||
2966 GGACTTAAAGGATGCTTCTCTGCTGAGATTACACCCACTAGCCAC 3015
284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
|||||
3016 CACTTTTTCCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCGGCAG 3065
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317
|||||
3066 CTCACCTGGACCCGACTGCCCAAGGGTTCAAGAACTCCCGACCATCTT 3115
317 eaSpGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
|||||
3116 TGACGAAGCCCTACACAGAGACCTGGCCAACTTCAGGATCCACACCTC 3165
334 lnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350
|||||
3166 AGGTGACCTCTCTCCAGTACGTGGATGACCTGCTTCTGGCGGAGCCACC 3215
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuLeuSerAs 367
|||||
3216 AAACAGGACTGCTTAGAAGGACGAAAGGCCTACTGCTGGAATTGCTCA 3265
367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
|||||
3266 CCTAGGCTACAGAGCTCTGTGTAAGAGCCAGATTTGCAGGAGAGAGG 3315
384 alThrTyrLeuGlyTyrSerLeuArgGlyGlnArgTrpLeuThrGlu 400
|||||
3316 TAACATACTTGGGTACAGTTTACGGGACGGCAGCGATGGCTGACGGAG 3365
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG 417
|||||
3366 GCACCGAAGAAACCTGTATGCCAGATACCGGCCCAACACACGCAACAA 3415
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
|||||
3416 AATGAGAGATTTTGGGACAGCTGGATTTGCAGACTGTGGATCCGG 3465
434 lyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450
|||||
3466 GGTTCGACCTTAGCAGCCCACTCTACCCGCTAACCCAAAGAAAAGGG 3515
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467
|||||
3516 GAATTCCTCTGGGTCTCGAGCACCAAGGCATTTGATGCTATCAAAA 3565
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484
|||||
3566 GGCCTCTGCTGAGCGACCTGCTGCGCCCTCCCTGACGTAACTAACCT 3615
484 heThrLeuTyrValAspGluArgLysGlyValAlaAlaArgGlyValLeuThr 500
|||||
3616 TTACCTTTATGTGATGAGCGTAAGCGGTAGCCCGGGAGTTTAAACC 3665
501 GlnThrLeuGlyProTyrArgArgProValAlaTyrLeuSerLysLysLe 517
|||||
3666 CAACCTTAGGACCATGGAGAGACCTGTGCGCTACCTGTCAAAGAAGCT 3715
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaIaV 534
|||||
3716 CGATCTCTGACCCAGTGTGGCCCATATGCTCAAGGCTATCCAGCGTG 3765
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGln.AsnI 550
|||||
3766 TGGCATACTGGTCAAGGACGCTGACAAATTGACTTTGGGACAAGAAAT 3815

```

550 eThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProA 567
|||||
3816 AACTGTAATAGCCCCCATGATGGAGAACATGCTCGGAGCCCCAG 3865
|||||
567 sPArgTrpMetThrasnAlaArgMetThrHisTyrGlnSerLeuLeu 583
|||||
3866 ACCGATGATGACCAACGCCGATGACCCACTATCAAAAGCCTGCTTCTC 3915
|||||
584 ThrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLe 600
|||||
3916 ACAGAGGGGTACGCTCGCTCCACCAACCGCTCTCAACCCCTGCCACTCT 3965
|||||
600 uLeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuL 617
|||||
3966 TCTGCTCAAGAGACTGATGAACAGTACATCATGATGGCCATCACTAT 4015
|||||
617 euIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThr 633
|||||
4016 TGATTCAGAGACTGGGTCCGCAAGGACCTTACAGACATACCGCTGACT 4065
|||||
634 GlyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluL 650
|||||
4066 GGAGAAGTGTAACTGGTTCACCGGAGACGCTATGTGGTGGAAAG 4115
|||||
650 yLysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpA 667
|||||
4116 TAAGAGGATGGTGGGCGCGGTGGTGACGGGACCGCGCATCTGGG 4165
|||||
667 laSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAla 683
|||||
4166 CCAGAGCGCTCGCGGAGGAGACTTCAGACACAAAGGCTGAGCTCATGGCC 4215
|||||
684 LeuThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrTh 700
|||||
4216 CTCACGCAAGCTTCGGCTGCGCGAGGGAATCCATAAACATTTATAC 4265
|||||
700 rAspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrL 717
|||||
4266 GGACAGCAGGTATGCTTTGGCAGTGCACAGCTACATGGGCGCATCTATA 4315
|||||
717 yGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGlu 733
|||||
4316 AACAAAGGGGTGGCTTACCTCAGCAGGGGAGGAAATAAGAACAAAGAG 4365
|||||
734 GluIleLeuSerLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750
|||||
4366 GAAATCTTAAGCCTATTAGAGCCGTACATTTACCAAAAGGCTAGCTAT 4415
|||||
750 eIleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyA 767
|||||
4416 TATACACTCTCTGGACATCAGAAAGCTAAAGATCTCATATCCAGAGGAA 4465
|||||
767 sGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeu 783
|||||
4466 ACCAGATGCTGACCGGTGGCAAGCAGGCGCCAGCCAGGGGTAAACCTT 4515
|||||
784 LeuProIleIleGluThrProLysAlaProGluProArgArgGlnTyrTh 800
|||||
4516 CTGCTTATATAGAATGCCCAAGCCCGAGAACCCAGACGACAGTACAC 4565
|||||
800 rLeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrP 817
|||||
4566 CCTAGAAGACTGGCAAGAGATAAAAAGATAGACCA.TTCTCTGAGACTC 4614
|||||
817 roGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLys 833
|||||
4615 CGGAAGGGACCTGCTATACCTCAGATGGGAAGAAATCCTGCCCCACAAA 4664
|||||
834 GluGlyLeuGluTyrValGlnGln.IleHisArgLeuThrHisLeuGlyT 850
|||||
4665 GAAGGTTAGATATGTCACCAAGATACATGCTTAACCCACCTAGGAA 4714
|||||

850 hrLysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArg 866
|||||
4715 CTAACACCTGCAGCAGTGGTTCAGAACATCCCTTATCATGTTCAGG 4764
|||||
867 LeuProGlyValAlaAspSerValValLysHisCysValProCysGlnLe 883
|||||
4765 CTACCAGAGTGGCTGACTCGGTGGTCAACACATGTGTGCTCCCTGCCAGCT 4814
|||||
883 uValAsnAlaAsnProSerArgIleProGlyLysArgLeuArgGlyLys 900
|||||
4815 GGTATATGTAATCCTTCAGATGCTCCAGGAAGAGACTAAGGGGAA 4864
|||||
900 erHisProGlyAlaHisTrpGluValAspPheThrGluValLysProAla 916
|||||
4865 GCCACCCAGCGCTCACTGGAAAGTGGACTTCACTGAGGTAAAGCGCGCT 4914
|||||
917 LysTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTr 933
|||||
4915 AAATATGGAAACAAATACCTATTGGTTTTGTAGACACCTTTTCAGAGTG 4964
|||||
933 pValGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysL 950
|||||
4965 GGTAGAGGCTTATCCTACTAAGAAAGAGACTTCAACCGTGGTAGCTAAA 5014
|||||
950 ystIleLeuGluGluIlePheProArgPheGlyIleProLysValIleGly 966
|||||
5015 AAATACTGGAAGAAATTTTCCAAGATTTGGAATACCTAAGGTAATAGG 5064
|||||
967 SerAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLys 983
|||||
5065 TCAGACAAATGCTCCAGCTTTTGTGCCAGGTAACTCAGGACTGGCCAA 5114
|||||
983 sIleLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerS 1000
|||||
5115 GATATTGGGATGATTGGAACTGCATTGTGCATACAGACCCCAAGCT 5164
|||||
1000 erGlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLys 1016
|||||
5165 CAGGACAGGTAGAGAGGATGAATAGAACCATTAAGAGAGACCCCTTACTAA 5214
|||||
1017 LeuThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheVa 1033
|||||
5215 TTGACCGCGAGACTGGCGTTAATGATTGGATAGCTCTCTGCCCTTGT 5264
|||||
1033 lLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrG 1050
|||||
5265 GCITTTTAGGGTTAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATG 5314
|||||
1050 luLeuLeuTyrGlyGlyProProLeuAlaGluIleAlaPheAlaHis 1066
|||||
5315 AATTACTCTACGGGGGACCCCCCATTTGGTAGAATTTGCTCTGTACAT 5364
|||||
1067 SerAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLe 1083
|||||
5365 AGTGCTGATGTGCTGCTTCCAGCCCTTGTCTCTAGGCTCAAGGCACT 5414
|||||
1083 uGluTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSer. 1099
|||||
5415 TGAGTGGGTGAGACACAGCGGTGGAGCAACTCCGGGAGGCGCTACTCAG 5464
|||||
1100 ..GlyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerVal 1115
|||||
5465 GAGGAGAGACTTGCAGATCCCATCGTTTCCAAGTGGGAGATTCAAGTC 5514
|||||
1116 TyrValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyPr 1132
|||||
5515 TACGTTAGACGCCACCGTGCAGGAAACCTCGAGACTCGGTGGAGGSCCC 5564
|||||
1132 oTyrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIleP 1149
|||||
5565 TTATCTGCTACTTTTGACCACACACCGCTGTGAAGTCCAAAGTCAAGGAATCT 5614
|||||
1149 roThrTrpIleHisAlaSerHisValLysProAlaProProPheAspSer 1165
|||||

|||||
 5615 CCACCTGGATCCATCCACGTTAAGCGCGGCCACCTCCCGATTCCG 5664
 1166 GlyTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisAr 1182
 |||||
 5665 GGGTGGAAAGCGAAAGAGACTGAAATCCCGTTAAGCTTCGCTCCATCG 5714
 1182 gLeuValProTyrSerAsnAsn 1190
 |||||
 5715 CGTGGTTCCTTACTCTGTCATAAC 5739

seq_name: gb_pat:AX052638

seq_documentation_block: 7873 bp DNA PAT 12-JAN-2001
 LOCUS AX052638
 DEFINITION Sequence 23 from Patent WO0071726.
 ACCESSION AX052638
 VERSION AX052638.1 GI:12226828

KEYWORDS
 SOURCE porcine endogenous retrovirus.
 ORGANISM porcine endogenous retrovirus
 Viruses; Retroviral viruses; Retroviridae; Mammalian type C
 retroviruses; 1-Mammalian type C virus group.

REFERENCE
 AUTHORS Federspiel M.J.
 TITLE Methods to inhibit infectious agent transmission during
 xenotransplantation
 JOURNAL Patent: WO 0071726-A 23 30-NOV-2000;
 Mayo Medical Ventures (US)

FEATURES
 source
 1. .7873
 /organism="porcine endogenous retrovirus"
 /db_xref="taxon:61673"
 BASE COUNT 2200 a 1914 g 1798 t
 ORIGIN

alignment_scores:
 Quality: 5862.00 Length: 1194
 Ratio: 5.106 Gaps: 4
 Percent Similarity: 96.147 Percent Identity: 94.137

alignment_block:

US-09-171-553B-5 x AX052638 ..

Align seg 1/1 to: AX052638 from: 1 to: 7873

1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17
 1637 GGAAGAGCGGGTTCGGACCCCTCCCGAGCGCCAGGGTAACCTTGAAGT 1686
 17 iGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34
 1687 GGAGGGGCACACAGTTGAGTTCCTGGTTGATACCGGAGCGCAACATTGAG 1736
 34 allLeuLeuGlnProLeuGlyLysLeuLysGlyLysSerTrpValMet 50
 1737 TGCTACTACAGCCATTAGGAAACTAAAGATAAAAAATCCTGGGTGATG 1786
 51 GlyAlaThrGlyGlnArgGlnTyrProTyrProTyrThrArgArgThrValas 67
 1787 GGTGCCACAGGGCAACAACAGTAGTATCCATGGACTTACCCGAAGAACAGTTGA 1836
 67 pLeuGlyValiGlyArgValThrHisSerPheLeuValIleProGluCysP 84
 1837 CTGGGAGTGGGAGCGGTAAACCCACTCGTTCTGTTATACCTAGTGGTGC 1886
 84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100
 1887 CAGCACCCCTCTTAGGTAGAGACTTATTGACCAAGATGGGAGCACAAT 1936
 101 SerPheGluGlnGlyArgProGluValSerValAsnLysProIleTh 117
 |||||

1937 TCTTTTGAACAAGGAAACACAGAGTGTCTGCAATAACACAACTATCAC 1986
 117 rValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGln 134
 |||||
 1987 TGTGTTTACCCCTCCAATTAGATGACGAATATGACTATATCTCTCCCTAG 2036
 134 allysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150
 |||||
 2037 TAAAGCCTGATCAAAATATACAAATCTGGTTGGACAGTTTCCCAAGCC 2086
 151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnVa 167
 |||||
 2087 TGGGCAGAAACCGCAGGATGGTTTGGCAAGCAAGTTCCCCACAAGT 2136
 167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184
 |||||
 2137 TATTCAACTGAAGGCCAGTGCACACCACTGTCAGTCAGACAGTAGTCCCT 2186
 184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200
 |||||
 2187 TGAGTAAAGAGCTCAAGAGGAATTCGGCCGATGTCCAAAGATTAATC 2236
 201 GlnGlnIlyleLeuValProValGlnSerProTyrProHisProLeuLe 217
 |||||
 2237 CAACAGGCGATCCTAGTTCCTGTCCAATCTCCCTGGAATACCTCCCTGCT 2286
 217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234
 |||||
 2287 ACCGGTTAGAAAGCCCTGGGACTTAATGACTATCGACAGTACAGGACTTGA 2336
 234 rGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250
 |||||
 2337 GAGAGTCAATAAAGGTGCGAGGATATACCCCAACAGTCCCGAACCCCT 2386
 251 TyrAsnLeuLeuSerAlaLeuProGluArgAsnTrpTyrThrValle 267
 |||||
 2387 TATAACCTCTTGTGTCTCTCCACCCCAACGAGCTGTTATACAGTATT 2436
 267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284
 |||||
 2437 GGACTTAAAGGATGCTTTTCTGCTGAGATTACACCCCACTAGCCCAAC 2486
 284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
 |||||
 2487 CACTTTTTCCTTCCAAATGGAGATCCAGGTACGGGAAGAACCGGGCAG 2536
 301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317
 |||||
 2537 CTCACCTGGACCCGACTGCCCAAGGGTTCAAGAACTCCCGACCATCTT 2586
 317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
 |||||
 2587 TGACGAAGCCCTTACACAGAGACTGCGCAACTTCAGGATCCCAACACCTC 2636
 334 InValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350
 |||||
 2637 AGGTGACCTCTCCAGTACGTGGATGACCTGCTTCTGGCGGGAGCCAC 2686
 351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuGluLeuSerAs 367
 |||||
 2687 AAACAGGACTGCTTAGAAAGCAGCAAGGACACTTACTGCTGGAATTGCTGA 2736
 367 pleuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
 |||||
 2737 CTTAGGCTACAGAGCTCTGTTAAGAAAGCCAGATTTTCAGAGAGAGAG 2786
 384 alThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400
 |||||
 2787 TAACATACTTGGGTACAGTTCGGGGACGGCGACGATGCTGACGAG 2836
 401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysGl 417
 |||||
 2837 GCACGGAAGAAACTGTAGTCCAGATACCGGCCCCCAACACAGCCACAAC 2886

417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
 |||||
 2887 AGTGAGAGAGTTTGGGACAGCTGGATTTCAGACTGTGGATCCCGG 2936
 |||||
 434 lypPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGly 450
 |||||
 2937 GGTTCGCGACCTTAGCAGCCCTCTACCCACTAACCAAGAAAGGG 2986
 |||||
 451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLY 467
 |||||
 2987 GAATTCCTCGGCTCCCTGAGCACCAGAGGCAATTGATGCTATCAAAA 3036
 |||||
 467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484
 |||||
 3037 GGCCCTGCTGAGCGCACCTGCTCTGGCCCTCCCTGACGTAAACCCCT 3086
 |||||
 484 heThrLeuTyrValAspGluArgLysGlyValAlaAlaArgGlyValLeuThr 500
 |||||
 3087 TTACCCCTTTATGTGATGAGCGTAAGGGAGTAGCCCGGGGAGTTTAAACC 3136
 |||||
 501 GlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLY 517
 |||||
 3137 CAAACTCTAGGACCATGGAGGAGACTGTGCTACCTACCTGTCAAAGAAGCT 3186
 |||||
 517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaLav 534
 |||||
 3187 CGATCCTGTAGCCAGTGGTGGCCCGTATGCCCTGAAGGCTATCGCAGCTG 3236
 |||||
 534 aAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550
 |||||
 3237 TGCCCATACTGTCAAGGAGCGTGCAAAATTGACTTTGGGACAGAATATA 3286
 |||||
 551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567
 |||||
 3287 ACTGTAATAGCCCCCATCGTTGGAGAACATCGTTTCGGACGCCCCAGAG 3336
 |||||
 567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeu 584
 |||||
 3337 CCGATGGATGACCAAGCCCGCATGCCACTATCAAAAGCCTGCTCTCA 3386
 |||||
 584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600
 |||||
 3387 CAGAGAGGTTCAGTTCGCTCCACGAGCGCTCTCAACCCCTGCCACTCTT 3436
 |||||
 601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617
 |||||
 3437 CTGCCCTCAAGACACTGATGACACCTGACTCATGATTGCCATCAACTATT 3486
 |||||
 617 ulleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634
 |||||
 3487 GATTGAGGAGACTGGGGTCCGCAAGACCTTACAGACATACCGCTGACTG 3536
 |||||
 634 lylGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650
 |||||
 3537 GAGAACTGTAACTCGTTCACTGACGGAAGCAGCTATGATGGAAGGT 3586
 |||||
 651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667
 |||||
 3587 AAGAGATGGCTGGCGCGCGGTGGTGACGGGACCCGCGCATCTGGGC 3636
 |||||
 667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684
 |||||
 3637 CAGCAGCCTGCCGGAAGAACTTCAGCACAAAGGCTGAGCTCATGGGCC 3686
 |||||
 684 euThrGlnAlaLeuArgLeuAlaGluLysSerIleAsnIleTyrThr 700
 |||||
 3687 TCACGCAAGCTTTGGCGCTGGCGGAAGGAAATCCATAACATTTATACA 3736
 |||||
 701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717
 |||||
 3737 GACAGCAGGTATGCCCTTTCGCGACTGCACACGTACACGGGGCCCATCTATAA 3786
 |||||

717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734
 |||||
 3787 GCAAGGGGGTTCCTTACCTCAGCAGGGAGGAAATAAAGAACAAAGAGG 3836
 |||||
 734 luleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750
 |||||
 3837 AAATTTCAAGCCTATTAGAAGCCTTACATTTGCCAAAAGGCTAGCTATT 3886
 |||||
 751 lIleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgLys 767
 |||||
 3887 ATACACTGTCTCGACATCAGAAAGCCAAAGATCCCATATCCAGAGGAA 3936
 |||||
 767 nGlnMetalAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784
 |||||
 3937 CCAGATGGCTGACCGGGTTGCCAAGCAGGAGCCAGGGGTGTTAACTTC 3986
 |||||
 784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800
 |||||
 3987 TGCCTATGATAGAAACACCCAAAGCCCAAGACCCGCGACGACAGTACCC 4036
 |||||
 801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817
 |||||
 4037 CTAGACACTGGCAAGAGATAAAAGATAGACCAGTTCTCTGAGACTCC 4086
 |||||
 817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834
 |||||
 4087 GGAAGGGACCTGCTATACCTCAGATGGGAAGAAATCCTGCCCCACAAAG 4136
 |||||
 834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850
 |||||
 4137 AAGGTTAGATAATGTCCACAGATACATCTGTAAACCCACCTAGGAACT 4186
 |||||
 851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867
 |||||
 4187 AACACCTGCAGCAGTTGGTCAGAAACATCTCTTATCATGTCTTGAGGCT 4236
 |||||
 867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeu. 883
 |||||
 4237 ACCAGGAGTGGCTGATTTCGGTGTCAAAACACTGTGTCCTGCCAGCTGG 4286
 |||||
 884 ValAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgLysE 900
 |||||
 4287 GTA..... 4289
 |||||
 900 rHisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaL 917
 |||||
 4290AAGCCGCTA 4299
 |||||
 917 ySTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrp 933
 |||||
 4300 AATACGAAACAAATATCTATTTGTTTGTAGACACCTTTTCAGGATGG 4349
 |||||
 934 ValGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLy 950
 |||||
 4350 GTAGAGCTTATCTACTAAGAAAGAGACTTCAACCGTGGTGGCTTAGAA 4399
 |||||
 950 sIleLeuGlu.GluIlePheProArgPheGlyIleProLysValIleGly 966
 |||||
 4400 AATACTGGAGGAAATTTTCCAAAGATTGGAAATACCTAAGGTAAATAGG 4449
 |||||
 967 SerAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLy 983
 |||||
 4450 TCAGACAAATGTCAGCTTTTCGTTGCCAGGTAAGTACAGACCCCAAGCT 4499
 |||||
 983 sIleLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerS 1000
 |||||
 4500 GATATTGGGATTGATTGGAAACTGCTTGTGCATACAGACCCCAAGCT 4549
 |||||
 1000 erGlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLys 1016
 |||||
 4550 CAGCAGAGGTAGAGAGGATGAATAGAACCACTTAAAGAGACCCCTTACTAAA 4599
 |||||
 1017 LeuThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheVa 1033
 |||||

284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
|||||
3608 CACTTTTTCGCTTGAATGAGAGATCCAGGTACGGGAAGAACCGGCAG 3657
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317
|||||
3658 CTCACCTGGACCGGACTGCCCAAGGGTCAAGAACTCCCGGACCATCTT 3707
317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
|||||
3708 TGACCAAGCCCTACACAGAGACCTGGCCCAACTTCAGGATCCAAACCCCTC 3757
334 InValThrLeuLeuGlnTyrValAspLeuLeuAlaGlyAlaThr 350
|||||
3758 AGGTGACCCCTCCCTCAGTACGTGATGACCTGCTTCTGGCGGAGCCACC 3807
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAs 367
|||||
3808 AAACAGGACTGTTAGAAGGCACGAAGGCACCTACTGCTGGAATTTGCTGA 3857
367 pleuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
|||||
3858 CCTAGGCTACAGAGCCTCTGCTAGAAGGCCCGGACGATTTGCGAGGAGAGG 3907
384 alThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400
|||||
3908 TAAACATACCTGGGTACAGTTTGGGGAGCGGACGCGATGCTGACGGAG 3957
401 AlaArgLysLysThrValValGlnIleProAlaProThrAlaLysG1 417
|||||
3958 GCACGGAAAGAACTGTAGTCCAGATACCGGCCCGCCCAACACAGCCAAACA 4007
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
|||||
4008 AGTGAGAGAAATTTTGGGGACAGCTGGATTTTGAGACTGTGGATCCGG 4057
434 lyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450
|||||
4058 GGTTTTGGCAGCTTAGCAGCCCACTCTACCCACTAACCCAAAGAAAAGGG 4107
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467
|||||
4108 GAATTTCTCTGGGCTCCCTGAGCACCAGAGGCAATTTGATGCTATCAAAA 4157
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484
|||||
4158 GGCCTGTGTGACGCGACCTGCTCTGGCCCTCCCTGACGTAACCTAAACCT 4207
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500
|||||
4208 TTACCCCTTATGTGGATGAGCGTAAGGGAGTAGCCCGGGAGTTTAAACC 4257
501 GlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLe 517
|||||
4258 CAAACTCTAGGACCATGAGGAGACCTGTGCTACCTACCTGCTCAAGAAGCT 4307
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaVal 534
|||||
4308 CGATCCGTAGCCAGTGGTTGGCCCGTATGCTGAAGGCTATCGCAGCTG 4357
534 aAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550
|||||
4358 TGGCCCATCTGCTCAAGGACGCTGCAAAATTTGACTTTGGGACAGAAATA 4407
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567
|||||
4408 ACTGTAATAGCCCCCCTGCTGGAGAACATCGTTTCGAGCCCCCAGA 4457
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuT 584
|||||
4458 CCGATGGATGACCAAGCCCGCATACCCACTATCAAGCCCTGCTTCTCA 4507
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600
|||||
4508 CAGAGAGGGTCACTGCTCGTCCACACCGCGCTCAACCTCCGCACTCTT 4557
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617
|||||
4558 CTCCTTGAAGAGACTCATGAACAGTGAATCATGATGGCATCACTATT 4607
617 urileGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634
|||||
4608 GATTGAGGAGACTGGGTCGCCAAGGACCTTACAGACATACCGCTGACTG 4657
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650
|||||
4658 GAGAAGTGTTAACCTGGTTCACTGACGGAAGCAGCTATGTGTGGAAGT 4707
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667
|||||
4708 AAGAGATGGCTGGGCGGCGGTGGTGGAGGAGCCCGCACGATCTGGGC 4757
667 aserSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684
|||||
4758 CAGCAGCCTGCCGGAAGAACTTCAGCACAAAAGGCTGAGCTCATGGCC 4807
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700
|||||
4808 TCACGGAAGCTTTGCGGCTGGCCGAAGGAAATCCATAACATTTATACA 4857
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717
|||||
4858 GACAGCAGTATGCTTTGCGACTGCACAGTACACGGGGCCATCTATAA 4907
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734
|||||
4908 GCAAAAGGGGTTGCTTACCTCAGCAGGGGGAATAAAGAACAAAGAGG 4957
734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750
|||||
4958 AAATTTAAAGCCCTATTAGAAGCCTTACATTTGCCAAAAGGCTAGCTATT 5007
751 ileHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767
|||||
5008 ATACACTGTCTTGACATCAGAAAGCAAGATCCCATATCCAGAGGGAA 5057
767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784
|||||
5058 CCAGATGGCTGACCGGTTGCCAAGCAGGCGCCAGGGTGTAAACCTTC 5107
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800
|||||
5108 TGCTATGATAGAAACACCCAAAGCCCGACAGACCCGAGCAGACACAC 5157
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817
|||||
5158 CTGAAGACTGGCAAGAGATAAAAAGATAGACAGTCTCTCTGAGACTCC 5207
817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834
|||||
5208 GGAAGGACCTGCTATACCTCAGATGGGAAGGAAATCCTGCCCCCAAG 5257
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850
|||||
5258 AAGGTTAGATATATGCCACAGATACATCGTCTAACCCACCTAGGAAC 5307
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValleuArgLe 867
|||||
5308 AAACACCTGCAGCAGTTGGTCAGAACATCTCTTATCATGTTCTGAGGCT 5357
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884
|||||
5358 ACCAGGAGTGGCTGATTGGTGGTCAAAACACTGTGTGCCCTGCCAGCTGG 5407
884 alAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900

|||||
5408 TTAATGCTAATCCTTCCAGATACTCCAGGAAGAGACTAAGGGGAAGC 5457
901 HisProGlyValHisTyrGluValAspPheThrGluValLysProAlaLy 917
|||||
5458 CACCAGCGCTCACTGGGAAGTGGACTTCACTGAGGTAAAGCCGGCTAA 5507
917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934
|||||
5508 ATACGGAACAATATCTATTGGTTTGTGTAGACACCTTTTCAGGATGG 5557
934 alGluAlaTyrProThrLysGluThrSerThrValValAlaLysLys 950
|||||
5558 TAGAGGCTTATCCTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAGAAA 5607
951 IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967
|||||
5608 ATACTGGAGGAATTTTCCAGATTTGGAATACCTAAGGTAATAGGGTC 5657
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984
|||||
5658 AGACAATGCTCCAGCTTTCGTTGCCAGGTAAGTCAGGGACTGGCCAAAGA 5707
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000
|||||
5708 TATTGGGGATTAAATTGGAAACTGCAATTGTGCATACAGACCCCAAGCTCA 5757
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017
|||||
5758 GGACAGGTAGAGAGATGAATAGAACCATTAAGAGAGACCCCTTACTAAAT 5807
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034
|||||
5808 GACCGCGAGACTGGCGTTAATGATTGGATAGCTCTCCTGCCCTTTGTGC 5857
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050
|||||
5858 TTTTAGGGTTAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATGAA 5907
1051 LeuLeuTyrGlyGlyProProLeuAlaGluIleAlaPheAlaHisSe 1067
|||||
5908 TTACTCTACGGGGACCCCCCATTTGGTAGAAATGCTTCGGTACATAG 5957
1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084
|||||
5958 TGCTGACGTGCTGCTTTCCAGCCTTTGTTCTTAGGCTCAAGGCACCTG 6007
1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSer... 1099
|||||
6008 AGTGGGTGACACAACGCGTGGAGCAACTCCGGAGGCGCTACTCAGGA 6057
1100 GlyGlyAspLeuGlnVal 1105
|||||
6058 GGAGGAGACTTGCAGATC 6075

OM of: US-09-171-553B-6 to: GenEmbl.* out_format : pfs
Date: Feb 24, 2002 10:15 AM
About: Results were produced by the GenCore software, version 4.5.
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-o=/cgn2.1/USPTO_spool/US09171553/runat_22022002.122825.11149/app_query.fasta_1.3312
-DB=GenEmbl -QWMT=fastap -SUFFIX=feb22std.rge -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000
-OCAPOP=4.500 -OCAPEXT=0.050 -YGAPOP=10.000 -YGAPEXT=0.500
-FCAPOP=6.000 -FCAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFWT=pfs
-NORM-ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09171553@cgn1.1.13817 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-171-553B-6
Query length: 656
Database: GenEmbl.*
Database sequences: 1472140
Database length: -341344837
Search time (sec): 5853.310000

score_list:

Seq	Strd	Orig	Zscore	EScore	Len	Documentation
gb_vl:AF014162	+ 3515.00	4141.42	2.1e-222	1971	!	AF014162 Porcine endogenous re
gb_pat:AF6553	+ 3515.00	4129.20	1.0e-221	8209	!	A6553 Sequence 3 from Patent
gb_pat:AF6552	+ 3509.00	4122.13	2.5e-221	8196	!	A6552 Sequence 2 from Patent
gb_pat:AF6559	+ 3455.00	4070.56	1.9e-218	1974	!	A6559 Sequence 9 from Patent
gb_pat:AX002804	+ 3433.00	4039.72	9.7e-217	3482	!	AX002804 Sequence 3 from Patent
gb_vl:PEREN2	+ 3433.00	4039.72	9.7e-217	3482	!	Y12239 Porcine endogenous re
gb_vl:PER293657	+ 3421.00	4017.65	1.7e-215	8763	!	Y12239 Porcine endogenous re
gb_vl:PER17013	+ 3417.00	4013.91	2.7e-215	7808	!	Y17013 Porcine endogenous re
gb_vl:PER13818	+ 3417.00	4012.94	3.0e-215	8750	!	AJ13818 Porcine endogenous re
gb_vl:PER288587	+ 3414.00	3909.23	4.9e-219	8918	!	AJ13818 Porcine endogenous re
gb_vl:PER288588	+ 3317.00	4007.50	2.3e-209	2000	!	AJ288588 Porcine endogenous re
gb_vl:PER288589	+ 3278.50	3873.99	1.7e-207	1968	!	AJ288589 Porcine endogenous re
gb_vl:PER288592	+ 3278.50	3861.45	8.3e-207	2000	!	AJ288592 Porcine endogenous re
gb_vl:PER288590	+ 3108.50	3661.33	1.2e-195	1996	!	AJ288590 Porcine endogenous re
gb_vl:PER288591	+ 3027.50	3565.66	2.5e-190	2002	!	AJ288591 Porcine endogenous re
gb_vl:PER288587	+ 3005.50	3539.69	6.9e-189	1999	!	AJ288587 Porcine endogenous re
gb_vl:PER288586	+ 2918.50	3436.93	3.7e-183	2008	!	AJ288586 Porcine endogenous re
gb_pat:AX002802	+ 2594.50	3052.61	9.3e-162	2462	!	AX002802 Sequence 1 from Patent
gb_vl:PEREN1	+ 2594.50	3052.61	9.3e-162	2462	!	Y12238 Porcine endogenous re
gb_vl:PER13817	+ 2590.50	3036.93	7.0e-161	8849	!	AJ13817 Porcine endogenous re
gb_vl:PER293656	+ 2541.50	2979.00	1.2e-157	8918	!	AJ293656 Porcine endogenous re
gb_pat:AX052637	+ 2509.50	2946.32	7.8e-156	4918	!	AX052637 Sequence 22 from Patent
gb_pat:AX052633	+ 2497.50	2939.94	1.8e-155	1980	!	AX052633 Sequence 18 from Patent
gb_vl:AF130444	+ 2497.50	2939.94	1.8e-155	1980	!	AF130444 Porcine endogenous re
gb_com:AF038601	+ 2497.50	2928.73	7.4e-155	7333	!	AF038601 Sus scrofa porcine en
gb_pat:AF130474	+ 2497.50	2928.73	7.4e-155	7333	!	AF130474 Sequence 2 from patent
gb_vl:PER288584	+ 2477.50	2916.26	3.7e-154	1996	!	AJ288584 Porcine endogenous re
gb_pat:AX052638	+ 2474.50	2900.96	2.6e-153	7873	!	AX052638 Sequence 23 from Patent
gb_vl:PER288585	+ 2465.50	2902.03	2.3e-153	2009	!	AJ288585 Porcine endogenous re
gb_com:AF296168	+ 2372.50	2792.97	2.7e-147	1840	!	AF296168 Sus scrofa porcine en
gb_pat:AF130473	+ 2360.00	2765.56	9.1e-146	8060	!	AF130473 Sequence 1 from patent
gb_com:AF038599	+ 2360.00	2765.48	9.2e-146	8132	!	AF038599 Sus scrofa porcine en
gb_com:AF038600	+ 2353.00	2757.22	2.7e-145	8132	!	AF038600 Sus scrofa porcine en
gb_pat:AF130475	+ 2353.00	2757.22	2.7e-145	8132	!	AF130475 Sequence 3 from patent
gb_hlg:AC079471	- 1678.50	1935.35	1.6e-99	158455	!	AC079471 Mus musculus clone
gb_vl:AF055062	+ 1532.00	1799.62	5.8e-92	2046	!	AF055062 Gibbon leukemia virus
gb_vl:AF055061	+ 1530.00	1797.21	7.9e-92	2058	!	AF055061 Gibbon leukemia virus
gb_vl:AF055063	+ 1527.00	1793.86	1.2e-91	2013	!	AF055063 Gibbon leukemia virus
gb_vl:AF055060	+ 1522.00	1787.76	2.7e-91	2058	!	AF055060 Gibbon leukemia virus
gb_com:AF151794	+ 1516.50	1769.19	2.9e-90	8431	!	AF151794 Phascolarctos cinereus

gb_pat:A60211 + 1508.50 1762.22 7.0e-90 6312 ! A60211 Sequence 7 from Pa
gb_pat:AR122287 + 1508.50 1762.22 7.0e-90 6312 ! AR122287 Sequence 7 from
gb_vl:GLU060065 + 1508.50 1759.80 9.6e-90 8379 ! U60065 Gibbon leukemia vi
gb_vl:AF055064 + 1500.00 1762.00 7.2e-90 2007 ! AF055064 Simian sarcoma-a
gb_hlg:AC079555 + 1494.00 1718.27 2.0e-87 144803 ! AC079555 Mus musculus c
seq_name: gb_vl:AF014162
seq_documentation_block:
LOCUS AF014162 1971 bp mRNA VRL 01-JUN-1998
DEFINITION Porcine endogenous retrovirus envelope protein (env) mRNA, complete
cds.
ACCESSION AF014162
VERSION AF014162.1 GI:3169735
KEYWORDS
SOURCE
ORGANISM
porcine endogenous retrovirus.
porcine endogenous retrovirus
Viruses; Retrovirdae; Retroviridae; Mammalian type C
retroviruses; i-Mammalian type C virus group.
REFERENCE
1 (bases 1 to 1971)
Haworth,C., Galbraith,D.N., Lees,G.M. and Smith,K.T.
Porcine endogenous retrovirus (PoEV) Env sequence
JOURNAL
Unpublished
TITLE
2 (bases 1 to 1971)
Haworth,C., Galbraith,D.N., Lees,G.M. and Smith,K.T.
Direct Submission
JOURNAL
Submitted (16-JUL-1997) Q-One Biotech Ltd., Todd Campus, Acre Rd.,
Glasgow G20 0XA, UK
FEATURES
source
1..1971
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
/cell_line="derived from cell line PK15"
/note="PoEV"
1..1971
/gene="env"
1..1971
/gene="env"
/codon_start=1
/product="envelope protein"
/protein_id="AAC17948.1"
/db_xref="GI:3169736"
/translation="MHPTLSRHLPTRGGEKRLRIPLSFASIAWFLTLITPQASSK
RLIDSSNPHRLSLTLITIDPTGTVNTGSGVAPRGTMWPELHFLCLRLINPAVKSTP
PNLVSYFGYCCPCKEYKCGSGESFCRWSCTVNDKWPISIQDRVKFSFVN
SPGKYKMKILYKDKSCSPDLYLKISFETKTKYKSKVDKVELGNSFLLYGGGAG
STLIRLIERGTGTPPVAMGPKVLAEGGPPALEPPHNLVPOLTSLRPDLTOPPSNS
TTGLIPTNTPRNSPGVPYKTCQRILSLIQAFAINSTDPTDSSCWLCSSGGPPYYE
GMAKERFNVTKEHNRNCTWGSRNKLTLEVSCKGTCTGKAPPSHQHLCYVTVYEQA
SENYLVPGYNRWACNTGLTPCVSTVFNQSKDFCVNQIVPRVYVHPHEVILDEYD
YRYNRKREPVSLLAVMLGLTGAVGTGTAALITGPQOLEKGLGELHAAWTEDLRA
LKESVNLFEESLSISEVVLNRRGLDLFLREGGLCAALKECCCFYVDHSGAIRDMS
NKLKKLLRRRRERERADGWEFNPSRPNWMTLLSALTGLPLVLLLLLTGPGCLINR
FVAFVRRVSQAVLMVLURQQVQGLSOGETDL"

BASE COUNT 530 a 490 c 482 g 469 t
ORIGIN
alignment_scores:
Quality: 3515.00 Length: 656
Ratio: 5.358 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-171-553B-6 x AF014162 ..
Align seq 1/1 to: AF014162 from: 1 to: 1971
1 MetHisProThrLeuSerArgArgHisLeuProThrArgGlyGlyGluPr 17
| ATGTCATCCCACTTAAGCGGGCCACCTCCCGACCTCGGGGTGAGAGCC 50
17 oLysArgLeuArgileProLeuSerPheAlaSerIleAlaTrpPheLeuT 34

|||||
51 GAAAGACTGAGAAATCCCTTAAAGCTTCGCTCCATCGCTGTTCTTCTTA 100
34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50
|||||
101 CTCTAAACAATACTCCCGAGGCGAGTAGTAAAGCCCTTATAGACAGCTCG 150
51 AspProHisArgProLeuSerLeuThrTrpLeuIleAspProAspTh 67
151 AACCCCATAGACCTTTATCCCTTACCTGGCTGATTATTAGCCCTGATAC 200
67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrpT 84
|||||
201 GGGTGTCACTGTAATAGACCTCGAGGTGTCTCTAGAGGACCTGGT 250
84 rpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
251 GGCCTGAACCTGCATTTCTCCCTCGGATTGATTAAACCCCGCTGTTAAAGC 300
101 ThrProAsnLeuValArgSerTyrGlyPheTyrCysProGlyTh 117
301 ACACCTCCCAACCTAGTCCGTAGTTATGGGTTCCTATTGCTGCCAGGCAC 350
117 rGluLysGluLysTyrCysGlyGlySerGlyGluSerPheCysArgAtgt 134
351 AGAGAAGAGAATACTGTGGGGTCTCGGGAATCCCTCTGTAGAGAT 400
134 rpSerCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150
401 GGAGCTGCCTCACCTCCAACGATGGAGACTGGAAATGCCGATCTCTCTC 450
151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrIly 167
451 CAGGACCGGGTAAAATTCCTTTGTCAATTCGCGCCGGGCAAGTACAA 500
167 smMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAspT 184
501 AATGATGAACATATATAAGATAAAGCTGCTCCCATCAGACTTAGATT 550
184 yrLeuLysIleSerPheThrGluArgLysThrGlyLysTyrSerLysVal 200
551 ATCTAAAGATAAGTTTCACTGAAAGGAAAAACAGGAAAAATATTCAAAAAGTG 600
201 AspLysTrpTyrGluLeuGlyAsnSerPheLeuLeuTyrGlyGlyAl 217
601 GATAAATGGTATGAGCTGGGGAATAGTTTTTATTATATGCGGGGGAGC 650
217 agLysThrLeuThrIleArgLeuArgIleGluThrGlyThrGluProP 234
651 AGGGTCCACTTTAACCATTCGCTTAGGNATAGAGCGGGGACAGAACCCC 700
234 roValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProProAla 250
701 CTGTGGCAATGGAGCCCGATAAGTACTGGCTGAACAGGGGGCCCCGGCC 750
251 LeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuArgPr 267
751 CTGGAGCCACCCATAACTTCGCGGTGCCCAATTAACCTCGCTCGCGCC 800
267 oAspIleThrGlnProProSerAsnSerThrThrGlyLeuIleProThrA 284
801 TGACATAACACAGCGCCCTAGGACACAGTACCCTGGATTGATTCCTACCA 850
284 snThrProArgAsnSerProGlyValProValLysThrGlyGlnArgLeu 300
851 ACACGCCTAGAAACTCCCGAGGTGTCTCTGTTAAGACAGGACAGAGACTC 900
301 PheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspProAs 317
901 TTCAGTCTCATCGGGAGCTTTTCAAGCCATCAACTCCACCGACCTGA 950
317 pAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTyrTyrG 334
|||||

951 TGGCACCTTCTTCTTGTGGCTTTGTCTATCTCTCAGGGCCTCCTTATTATG 1000
334 luGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArgAsn 350
1001 AGGGGATGGCTAAAGAAAGAAATTCATGTGACCAAGAGCATAGAAAT 1050
351 GlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSerGl 367
1051 CAATGTACATGGGGTCCGAAATTAAGCTTACCCCTCACTGAAGTTTCCGG 1100
367 yLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeuCyst 384
1101 GAAGGGGACATGCATAGGAAAGCTCCCATCCACCAACACACCTTTGCT 1150
384 yrSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeuValPro 400
1151 ATAGTACTGTGTTTATGAGCAGGCCCTCAGAAATCAGTATTTAGTACCT 1200
401 GlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValse 417
1201 GGTATTAACAGGTGGTGGCATGCAATACTGGGTAAACCCCTCTGTGTTT 1250
417 rThrSerValPheAsnGlnSerLysAspPheCysValMetValGlnIleV 434
1251 CACCTCAGTCTTCAACCAATCCAAAGATTTCTGTGTCTATGTCCTCAATCG 1300
434 alProArgValTyrTyrHisProGluGluValValLeuAspGluTyrAsp 450
1301 TCCCCCGAGTGTACTACCATCTCAGGAAGTGGTCTTGTATGAATATGAC 1350
451 TyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAlaVa 467
1351 TATCGGTATTAACCGACCAAAAGAGAACCCGCTATCCCTTACCCCTAGCTGT 1400
467 lMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAlaAlaL 484
1401 AATGCTCGGATTAGGACGGCGCTGGCTAGGAACAGGACAGCTGCC 1450
484 euIleThrGlyProGlnGlnLeuGlyLysGlyLeuGlyGluLeuHisAla 500
1451 TGATCACAGGACACAGCAGCTAGAGAAAGGACTTGGTGAGCTACATGCG 1500
501 AlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnLeuGl 517
1501 GCCATGACAGAAAGATCTCCGAGCCTTAAAGAGTCTGTTAGCAACCTAGA 1550
517 uGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyL 534
1551 AGAGTCCCTGACTTCTTTGTCTGAAGTGGTCTACAGAACCCGGAGGGAT 1600
534 euAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaAlaLeuLysGlu 550
1601 TAGATCTGCTGTTTCTTAAGAGAGGTGGTATTATGTCAGCCTTAAAAAGAA 1650
551 GluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMetAs 567
1651 GAATGTTGCTTCTATGTAGTACATCTCAGGAGCCATCAGACACTCCATGAA 1700
567 nLysLeuArgLysLysLeuGluArgArgArgGluArgGluAlaAspG 584
1701 CAAGCTTAGAAAAAAGTTAGAGGGCTCGAAGGAAAGAGAGCGTGACC 1750
584 lnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThrThrLeu 600
1751 AGGGGTGGTTTGAAGGATGGTTCAACAGGTCTCCTTGGATGACACCCCTG 1800
601 LeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeuThrVa 617
1801 CTTTCTCTCTGAGGGGGCCCTTAGTAGTCTCTCTCTCTGTTACTTACAGT 1850
617 lGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgVals 634
1851 TGGCCCTTGCTTAATTAATAGGTTTGTGGCTTGTGGCTTGTAGAGACGAGTGA 1900

```

634 erAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeuLeuSer 650
|||||
1901 GTGCAGTCCAGATGCTAGTACCTAGGCAACAGTACCAAGGCTTCTGAGC 1950

651 GlnGlyGluThrAspLeu 656
|||||
1951 CAAGGAGAACTGACCTC 1968

seq_name: gb_pat:A66553

seq_documentation_block:
LOCUS A66553 8209 bp DNA PAT 29-MAR-1999
DEFINITION Sequence 3 from Patent WO9740167.
ACCESSION A66553
VERSION A66553.1 GI:4538106
KEYWORDS
SOURCE
ORGANISM
porcine endogenous retrovirus.
Virus; Retrovirus; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
REFERENCE
1 (bases 1 to 8209)
AUTHORS Galbraith, D.N., Haworth, C., Lees, G.M. and Smith, K.T.
TITLE PORCINE RETROVIRUS
JOURNAL Patent: WO 9740167-A 3 30-OCT-1997;
Q ONE BIOTECH LTD (GB)
FEATURES
Location/Qualifiers
1..8209
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
588..2162
/codon_start=1
/product="VIRAL CORE PROTEIN"
/protein_id="CAB39347.1"
/db_xref="GI:4538107"
/translation="MGQIVTTPLSLTDLHWTEVRSRAHNLVSQVKKGPWQTFCASEWP
TFDQWPSSEGTNSBIIILAVKAIIFQTGPGSHDQPEYILTQDLAEDPPWPVKRWLNL
KPKPGPRILALGKKNHSAEKVEPSSYLPDRGAADLAGTPTCSPPLSTGCGCEG
TSAPGAPVEGPAAGTSSRGATPTERTDEIALLPLRTYGPMPGQLOLPQYWPFS
ADLYNKTNNHPFSEDORLTGLVESLMESHQPTWDDCOLOLTLFTTEERILLEA
RLNVCADGPTQLONEIDMGFPPLRPGWDYNTAGRESLKIYRQALVAGLQASLRD
TNLAVRVMQGNPEPPSVFLERLMEAFRRFTFPDPTSEAQKASVALAFICQASDIR
KKLRLGLEQAEALDLVREAEKVKYVRRTEEEKREKEKEEREERDRRQEKNL
KILAAVVEGKSSRERERFERKIRSGPRQSGNLNRTPLDKDQCAKCEKKGHWARNCPK
KGNKGPKVLALEEDK"
<2163..5747
/codon_start=1
/product="REVERSE TRANSCRIPTASE"
/protein_id="CAB39348.1"
/db_xref="GI:4538108"
/translation="GRRGSDPLPEPRVTILKVEGPVEFLVDTGAEHVSLLQPLCKLKE
KSWWVGATGQRYPTWTRRTVYDLGVNTHSFLVIPCEPVLGRLDRLTNKGQISF
EGREPVSNKAPITVILQLDDETRLSYQVKKPDQIQSWLEQPPQAWAETAGMGLA
KQVPQVTLKASATPSVVRQYPLSREAREGIWPHVORLIQOGLIVPVQSPWNTPLLP
VRKPGTNDYRPVQDLREYNKRVQDIHTPNPNNILSALPPERNWYTLDLKDAFCL
RLHPTSQLPFAEPWRDPTGRTQDITRLPGQFNKSPITFDEALHRLANFRJQHPO
VTLQVDDLLLAGATKODCLEGTAKLLESLDGLYRASAKKAOICREVTYLYSLR
GQORWLTARKTVVOIPAPTAKOVREFLGTAGFCRLWIPGFATLAPLYLPKXG
GFWAPAEHQKADA LKALLSAPALADPTKPTFLYVDERKGVARGVLTQTGLPWR
PVALYSLKLDPVQSWGVCPLKAIAVAAILVKDADKLTLGQNLITVIAHLENIVRQP
DRWNTNARMTHYVQSLLTERTVFAPALNATLLPEETDEPVTVDCHOLLIETGVR
KDLTDIPITGEVLTWFTDGSVSVVEGKMGAAVVDGTRTIWASSLPETGSAQAEML
ALTOALRLAEGKS INIYDTSRYAFATAHVGA IYKORGLLT SAGREIKNKEILSLE
ALHLEPKRLAIIHCPGHOKAKDLISRGQMADRVAKQAQAVNLLPIETPKAPPREQ
YILEBQKEIKKIDQFSETPGTCYTSYKLEILPHKEGLEIVQOIHRLTHLTKHLQQL
VRTSPYHLRLPGVADSVVKKHCPQOLYNANPRIPPGKRLRGSHGPAHWEDVTEV
PAKYGNKYLIVFDPTFGWVEAYPTKSTVAVKILIEEPFRFGIPKTVGSDNGPA
FVAQVSGQGLAKTLGIDWKLHCAYRPOSSGOVERNRTIKETLTKLTETGTNDGNMALL
PVLFRVNTPGQFGLTPYELLYGGPPPLAETAFASADVLLSLOPSRLKALEWVQ
RAWKQLREAYSGDLQVPHRFQVGSVYVRHRRAGNLETRNKGKLVLLTPTAVKVE
GIPTWISHASHPAPPPDPSGWRAEKSTENPLKRLHRLVFPYNNNSPGQ"
5620..7590
/codon_start=1

```

```

/product="VIRAL ENVELOPE PROTEIN"
/protein_id="CAB39349.1"
/db_xref="GI:4538109"
/translation="MHPTLSRRHLPTRGGEPRKLRIPLSFASIAWFLTLTITPOASSK
RLIDSSNPHRPLSLTLWLIIDPTGTVTNSTGVRAPRTWHPPELHFLCLINLPAVKSTP
PNLVSYGFYCCPTEKEKYGSGSEFCRWSCVTSNDGDKWPISLQDRVRFSEFVN
SGPKYKMKLYKDKSCPSDYLKISFTEKTKGYSKVDKWEYELGNSFLLYGGAG
STGLIRLRIETGTFPPVAMGDPKVLAEQGPALPEPHNLVPVQLTSLRPDITQPSNS
TTGLIPTNTPRNSPGVPVKTGORFLSLIOGAFOAINSTDPAATSSCWLCSSGPPYVE
GMAKERFNTVKEHNOCTGSRNKLTLTEVSGKGTGIGRAPPSHOHLICYSTVVEGA
SENQYLPVGYNRWACNIGLTPCYSTSVFNQSKDFCVWVQIVPVRYTHPREVLDEVD
YRINRKPRESLTLAVMLGIGTAGVGTGTAALITGPQQLERKGLGEUHAAMTEDLRA
LKESVNLLEESLTSLSVYLNQRRGLDLFLREGGLCAALKEECFCFYDHSAGIRDSM
NKLKRLEREREADQGWFEGRNRPWMTTLLSALTGPLVVLVLLLLLVGPCLINR
FVAFVRRVSAVQIMVLNRQOVLSSOGETDL"
BASE COUNT 2168 a 2064 c 2154 g 1823 t
ORIGIN

alignment_scores:
Quality: 3515.00 Length: 656
Ratio: 5.358 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-171-553B-6 x A66553
Align seg 1/1 to: A66553 from: 1 to: 8209

1 MetHisProThrLeuSerArgArgHisLeuProThrArgGlyGlyGluPr 17
|||||
5620 ATGCATCCACAGTTAAGCGCGGCCACCTCCCGACTCGGGTGGAGAGCC 5669

17 OlyArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuT 34
|||||
5670 GAAAGACTCAGAATCCCTTAAGCTTCGCCTCCATCGCTCGTTCCTTTA 5719

34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50
|||||
5720 CTCACATAACTCCCGAGGCCAGTAGTAAGCCCTTATAGACAGCTCG 5769

51 AsnProHisArgProLeuSerLeuThrTrpLeuIleAspProAspTh 67
|||||
5770 AACCCCATAGACCTTTATCCCTTACCTGGCTGATTATTGACCTGTATAC 5819

67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84
|||||
5820 GGGTGTCACTGTAATAATAGCACTCGAGGTGTTGCTCTAGAGCACCTGT 5869

84 rpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
|||||
5870 GGCCTGAATGCATTTCTCCCTCCGATTCATTAACCCCTGTAAAAGC 5919

101 ThrProProAsnLeuValArgSerTyrglyPheTyrcysCysProGly 117
|||||
5920 ACACCTCCCAACCTAGTCCGTAGTTATGGGTTCTATTGCTGCCAGCAC 5969

117 rGluLysGluLysTyrcysGlySerGlyGluSerPheCysArgArgT 134
|||||
5970 AGAAGAAGAGAATACTGGGGGTTCGCGGAATCCCTTCTGTAGAGAT 6019

134 rpSerCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150
|||||
6020 GGAGCTGCTCACCCTCCACGATGAGAGACTGGAAATGGCCGATCTCTC 6069

151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrl 167
|||||
6070 CAGGACCGGGTAAAAATTCCTTTTGTCATTCGCGCCCGGGCAAGTACAA 6119

167 sMetMetLysLeuTyrlLysAspLysSerCysSerProSerAspLeuAsp 184
|||||
6120 AATGATGAACATATATAAGAAAGAGCTGCTCCCATCAGACTTAGATT 6169

```

```
184 yrluLysIleSerPheThrGluArgLysThrGlyLysTyrSerLysVal 200
6170 ATCTAAAGATAAGTTTCACTGAAGAGAAACAGAGAAATATTCAAAAGTG 6219
201 AspLysTyrTyrGluLeuGlyAsnSerPheLeuLeuTyrGlyGlyAl 217
6220 GATAAATGATGAGCTGGGAAATAGTTTTTATTATATGCGGGGGAGC 6269
217 aGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGluPro 234
6270 AGGGTCCACTTTAAACCATTCGCCTTAGGATAGAGACGGGACAGAACCC 6319
234 roValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProAla 250
6320 CTGTGGCAATGGGACCCGATAAGTACTGGCTGAACAGGGGCCCGCCG 6369
251 LeuGluProPheHisAsnLeuProValProGlnLeuThrSerLeuArgPr 267
6370 CTGGAGCCACCGCATTAAGTTCGGGTGCCCAATTAACCTCGCTGGGCC 6419
267 oAspIleThrGlnProProSerAsnSerThrThrGlyLeuIleProThrA 284
6420 TGACATAACACAGCGCCCTAGCAACAGTACCACCTGGATTGATTCCTACCA 6469
284 snThrProArgAsnSerProGlyValProValLysThrGlyGlnArgLeu 300
6470 ACACGGCTAGAAGTCCCGAGGTGTTCCTGTTAAGACAGGACAGAGACTC 6519
301 PheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspProAs 317
6520 TTCAGTCTCATCCAGGAGCTTTCCCAAGCCATCACTCCACCGACCCCTGA 6569
317 pAlaThrSerSerCysTyrTrpLeuCysLeuSerSerGlyProProTyrTyrG 334
6570 TGCCACTCTCTCTGTGTGGCTTTGTCTATCTCTCAGGGCTCTCTATTATG 6619
334 lucIlyMetAlaLysGluArgLysPheAsnValThrLysGlnHisArgAsn 350
6620 AGGGATGGCTTAAGAAAGAAATTCATGTGACCAAGAGCATAGAAAT 6669
351 GlnCysThrTrpGlySerArgAsnLysLeuThrLeuGluValSerG1 367
6670 CAATGTACATGGGGTCCCGAAATTAAGCTTACCCCTCACTGAAGTTTCCGG 6719
367 yLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeuCyst 384
6720 GAAGGGAGCATCATAGGAAAGCTCCCCCATCCCAACCAACACCTTTGCT 6769
384 yrSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeuValPro 400
6770 ATAGTACTGTGGTTTATGAGCAGGGCTCAGAAAATCAGTATTAGTACCT 6819
401 GlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValSe 417
6820 GGTATACAGAGTGGTGGCATGCAATAGCTGGTTAAACCCCTGTGTTTC 6869
417 rThrSerValPheAsnGlnSerLysAspPheCysValMetValGlnIleV 434
6870 CACCTCAGTCTTCAACCAATCCAAAGATTTCTGTCTGTCATGGTCCAAATCG 6919
434 aProArgValTyrTyrHisProGluGluValValLeuAspGluTyrAsp 450
6920 TCCCCCGAGTGACTACCATCTCTGAGGAAGTGGTCTCTGATGAATATGAC 6969
451 TyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAlaVa 467
6970 TATCGGTATACCGACCAAAAGAGAACCCGTATCCCTTACCCTAGCTGT 7019
467 lMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAlaAlaL 484
7020 AATGTCGATTAGGACCGCGCTGGCGTAGGAACAGGACAGCTGCC 7069
484 euIleThrGlyProGlnGlnLeuGluLysGlyLeuGlyGlnLeuHisAla 500
```

```
7070 TGATCACAGGACCAACACACACTAGAGAAAGAGACTTGTGTGAGCTACATCG 7119
501 AlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnLeuG1 517
7120 GCCATGACAGAAAGATCCGAGCCTTAAGAGAGTCTGTTAGCAACCTAGA 7169
517 uGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyL 534
7170 AGATCCCTGACTTCTTTGTCTGAAGTGGTTTACAGAACCGGAGGGAT 7219
534 euAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaAlaLeuLysGlu 550
7220 TAGATCTGCTCTTCTTAAGAGAGGTGGTATGTGCGAGCCTTAAGAAGAA 7269
551 GluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMetAs 567
7270 GAATGTTGCTTCTATGTAGATCACTCAGGAGCCATCAGAGACTCCATGAA 7319
567 nLysLeuArgLysLysLeuGluArgArgArgGluArgGluAlaAspG 584
7320 CAAGCTTTAGAAAAAAGTTAGAGAGCGCTCGAAGGGAAGAGAGCTGACC 7369
584 lGlyTyrPheGluGlyTyrPheAsnArgSerProTyrMetThrThrLeu 600
7370 AGGGTGGTTTGAAGGATGGTTCAACAGGTCTCTCTGGATGACCACTG 7419
601 LeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeuThrVa 617
7420 CTTTCTCTCTGACGGGGCCCTAGTAGTCTGCTCTCTGTTACTTACAGT 7469
617 lGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgValS 634
7470 TGGCCCTTGCTTAAATTAATAGTGTGTGGCTTTGTTAGAGAACGAGTGA 7519
634 erAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeuLeuSer 650
7520 GTGCAGTCCAGATCATGGTACTTAGGCAACAGTACCAGGCGCTCTGAGC 7569
651 GlnGlyGluThrAspLeu 656
7570 CAAGGAGAACTGACCTC 7587
seq_name: gb_pat:A66552
seq_documentation_block:
LOCUS A66552 8196 bp DNA PAT 29-MAR-1999
DEFINITION Sequence 2 from Patent WO9740167.
ACCESSION A66552
VERSION A66552.1 GI:4538105
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
BASE COUNT 2165 a 2061 c 2146 g 1820 t 4 others
ORIGIN
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
Location/Qualifiers
1..8196

```

alignment_scores:
Quality: 3509.00 Length: 656
Ratio: 5.357 Gaps: 0
Percent Similarity: 99.848 Percent Identity: 99.848

alignment_block:

US-09-171-553B-6 x A66552

Align seg 1/1 to: A66552 from: 1 to: 8196

1 MetHisProThrLeuSerArgHisLeuProThrArgGlyGlyGluPr 17
|||||
5606 ATGCATCCACGTTAAGCYGGCGCCACCTCCCGACTCGGGGTGGAGACC 5655
17 oLysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuTr 34
|||||
5656 GAAAAGACTGAGATCCCTTAAGCTTCGCCTCCATCGCTGGTTCCTTA 5705
34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50
|||||
5706 CTCTAACAAATACTCCCGAGCCAGTAGTAAACGCCATTATAGACAGCTCG 5755
51 AsnProHisArgProLeuSerLeuThrTrpLeuIleIleAspProAspTh 67
|||||
5756 AACCCCATAGACCTTTATCCCTTACCTGGCTGATTAATGACCCCTGATAC 5805
67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrpT 84
|||||
5806 GGGTGTCACTGTAATAGCATCTCAGAGGTGCTCCTAGAGGCACTCGGT 5855
84 rpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
|||||
5856 GGCCTGAACCTCAATTCCTGCCCTCGGATGATTAAACCCCGCTGTTAARAGC 5905
101 ThrProProAsnLeuValArgSerTyrglyPheTyrcysCysProGlyTh 117
|||||
5906 ACACCTCCCAACCTAGTCCGTAGTATGGGTTCTATTGCTGCCCGGCGAC 5955
117 rGluLysGluLysTyrcysGlyGlySerGlyGluSerPheCysArgArgT 134
|||||
5956 AGAGNAAGAGAAATACTGTGGGGTTCGGGGAAATCTCTGTAGGAGAT 6005
134 rpSerCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150
|||||
6006 GGAGCTGGCTCACCTCCCAACGATGGAGACTGGAAATGGCCGATCTCTCTC 6055
151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrlY 167
|||||
6056 CAGGACCGGGTAAATAATCTCCTTTGTCAATTCGGCGCCGGGCAAGTACAA 6105
167 sMetMetLysLeuTyrlYsAspLysSerCysSerProSerAspLeuAspT 184
|||||
6106 AATGATGAACTATATAAGATAAGAGCTGCTCCCCATCAGACTTAGATT 6155
184 yrLeuLysIleSerPheThrGluArgLysThrGlyLysTyrlYsSerLysVal 200
|||||
6156 ATCTAAAGATAAGTTTCACTGAAAGGAAACAGCAAAATAATTCAAAAGTG 6205
201 AspLysTrpTyrlYLeuLeuGlyAsnSerPheLeuLeuTyrlYsGlyAl 217
|||||
6206 GATAAATGATGATGAGCTGGGAAATAGTTTTTATATATGCGGGGGAGC 6255
217 aGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGluProp 234
|||||
6256 AGGGTCCACTTAAACCATTCGCTTAGATAGACGGGGGACAGAACCC 6305
234 roValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProProAla 250
|||||
6306 CTGTGGCAATGGGACCGGATAAAGTACTGGCTGAACAGGGGCCCGCCGCC 6355
251 LeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuArgPr 267
|||||
6356 CTGGAGCCACCGCAATACTTCCCGGTGCCCAATAAACCTCGCTGCGGCC 6405
267 oAspIleThrGlnProProSerAsnSerThrThrGlyLeuIleProThrA 284
|||||
6406 TGACATACACAGCGCGCTAGCACACGTACCACCTGGATTGATTCCTACCA 6455

284 snThrProArgAsnSerProGlyValProValLysThrGlyGlnArgLeu 300
|||||
6456 ACACGCTAGAACTCCCGAGGTTCCTGTTAAGACAGGACAGAGACTC 6505
301 PheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspProAs 317
|||||
6506 TTCAGTCTCATCCAGGGAGCTTTCCAAAGCCATCAACTCCACCGACCCCTGA 6555
317 pAlaThrSerSerCysTrpLeuSerLeuSerSerGlyProProTyrlYrG 334
|||||
6556 TGGCACCTTCTCTGTGTGGCTTTGTCTATCTCAGGGCCCTCCTTATTATG 6605
334 luGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArgAsn 350
|||||
6606 AGGGATGGCTAAAGAAAGAAATTCATGTGACCAAAAGAGCATAGAAAT 6655
351 GlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSerGl 367
|||||
6656 CAATGTACATGGGGTCCCGAAATAAGCTTACCTCCTCACTCAAGTTTCCGG 6705
367 yLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeuCysT 384
|||||
6706 GAAGGGACATGCATAGGAAAGCTCCCATCCACCAACACCTTTGCT 6755
384 yrSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeuValPro 400
|||||
6756 ATAGTACTGTGTTTATGACAGCGCTCAGAAATCAGTATTATTAGTACCT 6805
401 GlyTyraAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValse 417
|||||
6806 GGTATTAACAGGTGGTGGCATCACTACTGGGTTAACCCCTGTGTTTC 6855
417 rThrSerValPheAsnGlnSerLysAspPheCysValMetValGlnIleV 434
|||||
6856 CACCTCAGTCTTCAACCAATCCAAAGATTCTGTGTCATGGTCCAAATCG 6905
434 alProArgValTyrTyrlYHisProGluGluValValLeuAspGluTyrlYrAsp 450
|||||
6906 TCCCGCGAGTGTACTACCATCTCTGAGGAAGTGTCTCTGTGAATATGAC 6955
451 TyrlYrArgTyrlYrAsnArgProLysArgGluProValSerLeuThrLeuAlaVa 467
|||||
6956 TATCGGTATAACCGACCAAAAGAGAACCGGTATCCCTTACCTAGCTGT 7005
467 lMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAlaAlaL 484
|||||
7006 AATGCTCGGATTAGGGACGCGCTTGGCGTAGGAACAGGGAGCTGCCCC 7055
484 euIleThrGlyProGlnGlnLeuGluLysGlyLeuGlyGluLeuHisAla 500
|||||
7056 TGATCAGAGCACAGCAGCTAGAGAAAGGACTTGGTGAGCTACATGCG 7105
501 AlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnLeuGl 517
|||||
7106 GCCATGACAGAGATCTCCGAGCTTTAAAGGAGTCTGTAGCAACCTAGA 7155
517 uGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyL 534
|||||
7156 AGAGTCCCTGACCTTCTTGTCTGAAGTGGTCTTACAGAACCGGAGGGGAT 7205
534 euAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaAlaLeuLysGlu 550
|||||
7206 TAGATCTGCTGTTTCTTAAGAGAAAGGTGGTATTGTGCAGCCTTAAAGAA 7255
551 GluCysCysPheTyrlYrValAspHisSerGlyAlaIleArgAspSerMetAs 567
|||||
7256 GAATGTGCTTCTATGTAGTCACTCAGGAGCCATCAGAGACTCCATGAA 7305
567 nLysLeuArgLysLysLeuGluArgArgArgGluArgGluAlaAspG 584
|||||
7306 CAAGCTTAAAAAAGTTAGAGAGCGCTCGAAGGGAAGAGAGGCTGACC 7355
584 lngGlyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThrLeu 600

|||||
 7356 AGGGGTGGTTTGAAGGATGTTCAACAGGTCCTCTGGATGACCAACCCCTG 7405
 601 LeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuThrVa 617
 7406 CTTTCGCTCTGACGGGGCCCTAGTAGTCCCTGCTCTTACTTACAGT 7455
 617 lGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgValS 634
 7456 TGGGCTTCTCTTAATTAATAGCTTTGTTGCCCTTTGTAGAGAACCAGTGA 7505
 634 exAlaValGlnIleMetValLeuArgGlnGlnIleThrGlnGlyLeuLeuSer 650
 7506 GTGCAGTCAGATCATGCTAGTACTTAGCAACAGTACCAAGGCCCTTCTGAGC 7555
 651 GlnGlyGluThrAspLeu 656
 7556 CAAGGAGAACTGACCTC 7573

seq_name: gb_pat: A66559

seq_documentation_block: 1974 bp DNA PAT 29-MAR-1999

LOCUS A66559 Sequence 9 from Patent WO9740167.

ACCESSION A66559

VERSION A66559.1 GI:4538112

KEYWORDS

SOURCE unidentified.

ORGANISM unidentified.

REFERENCE 1 (bases 1 to 1974)

AUTHORS Galbraith, D.N., Haworth, C., Lees, G.M. and Smith, K.T.

TITLE PORCINE RETROVIRUS

JOURNAL Patent: WO 9740167-A 9 30-OCT-1997;

Q ONE BIOTECH LTD (GB)

FEATURES Location/Qualifiers

1..1974

source /organism="unidentified"

/db_xref="taxon:32644"

BASE COUNT 534 a 489 c 483 g 468 t

ORIGIN

alignment_scores:

Quality: 3455.00 Length: 658

Ratio: 5.291 Gaps: 2

Percent Similarity: 99.240 Percent Identity: 99.088

alignment_block:

US-09-171-553B-6 x A66559 ..

Align seg 1/1 to: A66559 from: 1 to: 1974

1 MetHisProThrLeuSerArgHisLeuProThrArgGlyGluPr 17

1 ATCATCCCAAGTAAAGTGGGCGCCACCTCCCGACTCGGGGTGGAGAGCC 50

17 oLysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeu 34

51 GAAAGACTGAGAAATCCCTTAAGCTTCGCTCCATCGCTGGTTCCTTA 100

34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50

101 CTCTAAACAATACTCCCGAGGCGAGTAGTAAAGCCCTTATAGACAGCTCG 150

51 AsnProHisArgProLeuSerLeuThrTrpLeuIleIleAspProAspTh 67

151 AACCCCATAGACCTTTATCCCTACTCGCTGATTTATGACCCCTGATAC 200

67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84

201 GGGTGTCACCTGTAATAGCACTCGAGGTGTGTGCTCTAGAGGACCTGGT 250

84 rpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
 251 GGCCTGAACGCAATTCGCTCCGATTGATTAAACCCCGCTGTAAAGC 300
 101 ThrProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117
 301 ACACCTCCCAACCTAGTCCGTAGTTATGGGTTCATTTGCTGCCAGGCAC 350
 117 rGluLysGluLysTyrCysGlyGlySerGlyGluSerPheCysArgArg 134
 351 AGACAAAGAGAAATACTGTGGGGTCTCTGGGAATCCTCTCTGAGGAGAT 400
 134 rpSerCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150
 401 GGAGCTCGTCACTCCACGATGGAGACTGGAATGGCCGATCTCTCTC 450
 151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTrp 167
 451 CAGGACCGGGTAAATTTCTCTTTGTCAATTCGGCCCGGGGCAAGTACAA 500
 167 sMetMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAsp 184
 501 AATGATGAAACTATATAAGATAAGAGCTGCTCCCATCAGACTTAGATT 550
 184 yrLeuLysIleSerPheThrGlu..ArgLysThrGlyLysTyrSerLys 200
 551 ATCTAAAGATAAGTTTCACTGAAAGGAAACAGAGAAATATTTCAAAG 600
 200 aAspLysTrpTyrGluLeuGlyAsnSerPheLeu..LeuTyrGlyGly 216
 601 TGGATAAATGGTAAGAGCTGGGGAATAGTTTTTATAAATATGCGGGGG 650
 216 yAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGlu 233
 651 AGCAGGTCCACTTAACCTATCCCTTAGGATAGAGACGGGACAGAAC 700
 233 roProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyPro 249
 701 CCCCTGTGCAATGGAGCCGATAAAGTACTGGCTGAACAGGGGGCCCCG 750
 250 AlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeu 266
 751 CCCCTGGAGCCCGCATACCTTGGCGGTGCCCAATTAACCTCGCTGG 800
 266 gProAspIleThrGlnProProSerAsnSerThrThrGlyLeuIlePro 283
 801 CCCTGACATAACACAGCGCGCTAGCAACAGTACCCTGGATTGATTCTTA 850
 283 hrAsnThrProArgAsnSerProGlyValProValLysThrGlyGlnArg 299
 851 CCAACAGCCCTAGAAACTCCCGGTTCTCTTAAAGACAGGACAGAGA 900
 300 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspPr 316
 901 CTCCTCAGCTCATCCAGGGAGCTTTCCAAGCCATCAACTCCACCGACC 950
 316 oAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProTrpTr 333
 951 TGATGCCACTTCTCTTGGCTTGTCTATCTCTCAGGGCCCTCTTATT 1000
 333 yrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArg 349
 1001 ATGAGGGGATGGCTAAAGAAAGAAATTCATGTGACCAAGACATAGA 1050
 350 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSe 366
 1051 AATCAATGTACATGGGGTCCCGAAATAAGCTTACCCTCACTGAAGTTTC 1100
 366 rGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeu 383
 1101 CGGGAAGGGGACATGCATAGGAAAGCTCCCCCATCCCAACACCTTT 1150
 383 yTyrSerThrValValThrGluGlnAlaSerGluAsnGlnTyrLeuVal 399

```
|||||
1151 GCTATAGTACTGTGTTTATGAGCAGGCCTCAGAAATCACTATTACTA 1200
400 ProGlyTyrAsnArgTyrTrpAlaCysAsnThrGlyLeuThrProCysVa 416
|||||
1201 CCTGGTTATACAGGTGGTGGCATGCAATACTGGTTAAACCCCTGTGT 1250
416 lSerThrSerValPheAsnGlnSerLysAspPheCysValMetValGlnI 433
|||||
1251 TTCCACCTCAGTCTCAACCAATCCAAAGATTGTGTCTCATGTCCAAA 1300
433 leValProArgValTyrTyrHisProGluGluValValLeuAspGluTyr 449
|||||
1301 TCGTCCCGAGTGCTACTACCATCTGAGGAAGTGGTCTTGATGAATAT 1350
450 AspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAl 466
|||||
1351 GACTATCGGTATACCGACCAAAAGAGAACCCGTATCCCTTACCCTAGC 1400
466 aValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAlaA 483
|||||
1401 TGTAATGCTCGGATTAGGACGGCGGTGGCGTAGGAACAGGACAGCTG 1450
483 laLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuGlyGluLeuHis 499
|||||
1451 CCCTGATCACAGGACACAGCAGCTAGAGAAAGGACTTGGTGAGCTACAT 1500
500 AlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnLe 516
|||||
1501 CGGCCCATGACAGAAGATCTCCGAGCCTTAAAGGAGTCTGTAAACACT 1550
516 uGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgG 533
|||||
1551 AGAAGAGTCCCTGACTTCTTGTCTGAAGTGGTTCTACAGAACCGGAGG 1600
533 lyLeuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaLeuLys 549
1601 GATTAGATCTGCTCTTCTAAGAGAAGGTGGGTTATGTGAGCCTTAAAA 1650
550 GluGluCysCysPheTyrValAspHisSerGlyAlaIleArgaspSerMe 566
1651 GAAGAATGTCTCTATGATAGATCACTCAGGAGCCATCAGAGACTCCAT 1700
566 tAsnLysLeuArgLysLysLeuGluArgArgArgGluGluAlaA 583
1701 GAACAAGCTTAGAAAAAAGTTAGAGAGCGCTCGAAGGGAAGAGAGCTG 1750
583 spGlnGlyTyrPheGluGlyTrpPheAsnArgSerProTyrMetThrThr 599
1751 ACCAGGGGTGGTTTGAAGGATGGTTCACAGGCTCCTTGGATGACCACC 1800
600 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuTh 616
1801 CTGCTTCTCTCTGACGGGCCCTAGTAGTCTGCTCTCTGTACTTAC 1850
616 rValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgV 633
1851 AGTTGGCCCTTGCTTAATAATAGGTTTGTGCTTTGTAGAGAACGAG 1900
633 alSerAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeuLeu 649
1901 TGAGTCAGTCCAGATCATGGTACTTAGGCAACAGTACCAAGGCCCTTCTG 1950
650 SerGlnGlyGluThrAspLeu 656
|||||
1951 AGCCAAGGAGAACTGACCTC 1971
seq_name: gb_pat:AX002804
seq_documentation_block:
LOCUS AX002804 3482 bp DNA PAT 21-AUG-2000
DEFINITION Sequence 3 from Patent WO9853104.
ACCESSION AX002804
```

```
VERSION AX002804.1 GI:9885132
KEYWORDS
SOURCE
ORGANISM porcine endogenous retrovirus.
porcine endogenous retrovirus
Viruses; Retroviral viruses; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
REFERENCE 1 (bases 1 to 3482)
AUTHORS Stoyer,J.P. and Weiss,R.A.
TITLE Detection of retroviral subtypes based upon envelope specific
sequences
JOURNAL Patent: WO 9853104-A 3 26-NOV-1998;
MEDICAL RES COUNCIL (GB); STOVE JONATHAN PAUL (GB)
FEATURES
source
1..3482
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
BASE COUNT 927 a 854 c 867 g 834 t
ORIGIN
```

alignment_scores:

Quality: 3433.00 Length: 658
Ratio: 5.257 Gaps: 2
Percent Similarity: 99.240 Percent Identity: 98.176

alignment_block:

US-09-171-553B-6 x AX002804 ..

Align seg 1/1 to: AX002804 from: 1 to: 3482

```
1 MetHisProThrLeuSerArgArgHisLeuProThrArgGlyGluPr 17
|||||
911 ATCATCCCCACGTTAGCTGGCCACCTCCCGACTCGGGGGGAGAGCC 960
|||||
17 oLysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuT 34
|||||
961 GAAAGACTCAGAAATCCCTTAAGCTTCGCTCCATCGCTGTTCTTFA 1010
|||||
34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuLeaspSerSer 50
|||||
1011 CTCTAACAAATACTCCCGAGGCAGTAGTAAACGCCTTATAGACAGCTCG 1060
|||||
51 AsnProHisArgProLeuSerLeuThrTrpLeuIleIleAspProAspTh 67
|||||
1061 AACCCCATAGACCTTTATCCCTTACCTGGCTGATTATTGACCTGATAC 1110
|||||
67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrpT 84
|||||
1111 GGGTGTCACTGTAATAGCACTCGAGGTTGTCTCTAGAGCAGCCTGT 1160
|||||
84 rpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
|||||
1161 GGCCTGAATGCTATTCTCCCTCCGATTGATTAAACCCCTGTTAAAGC 1210
|||||
101 ThrProProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117
|||||
1211 ACACCTCCCAACCTAGTCGTAGTTATGGTTCTATTGCTGCCAGGCAC 1260
|||||
117 rGluLysGluLysTyrCysGlyGlySerGlyGluSerPheCysArgArgT 134
|||||
1261 AGAGAAAGAGAAATACTGTGGGGTTCTGGGGAATCCTTCTTAGAGAT 1310
|||||
134 rpSerCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150
|||||
1311 GGAGCTGCTCACCTCCCAACGATGGAGACTGGAAATGGCCGATCTCTTC 1360
|||||
151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrLy 167
|||||
1361 CAGGACCGGGTAAATTCCTCTTGTCAATCCCGCCCGGCACTACAA 1410
|||||
167 sMetMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAspT 184
|||||
1411 AGTGATGAACCTATATAAGATAAGAGCTGCTCCCATCAGACTTAGATT 1460
```

```

184 yfLeuLysIleSerPheThrGlu...ArgLysThrGlyLysTyrSerLysV 200
1461 ATCTAAGATAAGTTTCTACTGAAGAAGGAAACAGGAAATATTCAAAAG 1510
200 aAspLysTrpTyrGluLeuGlyAsnSerPheLeu...LeuTyrGlyGlyG 216
1511 TCGATAAATGGTATGAGCTGGGGAATAGTTTTTATATAAATATGGCGGGG 1560
216 yAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGluP 233
1561 ACAGGGTCCACTTTAAACCATTCGGCTTAGGATAGAGACGGGGACAGAAC 1610
233 roProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProPro 249
1611 CCCCTGTGGCAGTGGGACCCGATAAGTACTGGCTGAACAGGGGCCCGCC 1660
250 AlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuAr 266
1661 GGCCTGGAGCCACCCATAACTTGGCGGTGGCCCAATTAACCTCGCTGCG 1710
266 gProAspIleThrGlnProProSerAspSerThrThrGlyLeuIleProT 283
1711 GCCTGACATAACACAGCCCGCTAGCAACGGTACCACCTGGATTGATTCCTA 1760
283 hrAsnThrProArgAsnSerProGlyValProValLysThrGlyGlnArg 299
1761 CCACACGCCCTAGAAACATCCCCAGGTTCCTGTTAAGACAGGACAGAGA 1810
300 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspPr 316
1811 CACTTCAGTCTCATCCAGGAGCTTTCGAAGCCATCAACTCCACCGACCC 1860
316 oAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTyrT 333
1861 TGATGCCACTTCTCTGTGTGCTTGTCTATCCCTCAGGGCCCTCCTTATT 1910
333 yrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArg 349
1911 ATGAGGGGATGGCTAAAGAAGGAAATTCATGTGACCAAGAGCATAGA 1960
350 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuGluValSe 366
1961 AATCAATGTACATGGGGTCCCGAATAAGCTTACCCTCAGTGAAGTTTC 2010
366 rGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeuC 383
2011 CGGGAAGGGGACATGCATAGGAAGAGCTCCCCATCCACCAACACCTTT 2060
383 ysTyrSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeuVal 399
2061 GCTATAGTACTGTGGTTTATGAGAGCGCTCAGAAAATCAGTATTAGTA 2110
400 ProGlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysVa 416
2111 CTGGTTATAACAGGTGGTGGGCATGCAATCTGGGTTAACCCCTGTGT 2160
416 lSerThrSerValPheAsnGlnSerLysAspPheCysValMetValGlnI 433
2161 TTCCACCTCAGTCTCAACCAATCCNAAGATTTCTGTGTCATGGTCCAAA 2210
433 leValProArgValTyrTyrHisProGluGluValValLeuAspGluTyr 449
2211 TCGTCCCGGAGTGTACTACCATCTCTGAGGAAGTGGTCTTCGATGAATAT 2260
450 AspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAl 466
2261 GACTATCGGTATACCCACCAAAAGAGAACCCGGTATCCCTTACCCTAGC 2310
466 aValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAlaAla 483
2311 TGTAAATGCTCGGATTAGGACGCGCTTGGCGGTAGGAACAGGACAGCTG 2360

```

```

483 laLeuIleThrGlyProGlnGlnLeuGlyLysGlyLeuGlyGluLeuHis 499
2361 CCCTGATCAGAGCACACAGCAGCTAGAGAAAGGACTTGGTGAGCTACAT 2410
500 AlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnLe 516
2411 CGGCCATGACAGAAGATCTCCGAGCCTTAGAGAGTCTGTTAGCAACCT 2460
516 uGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgG 533
2461 AGAAGAGTCCCTGACTTCTTTGCTGAAGTGGTTCTACAGAACCAGGAGG 2510
533 lyLeuAspLeuLeuPheLeuArgGluGlyLeuGlyLysAlaAlaLeuLys 549
2511 GATTAGATCTGCTCTTTCTTAAGAGAAGTGGTTATGTGCAGCCTTAAAA 2560
550 GluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMe 566
2561 GAGAATGTTGCTTCTATGATGATCACTCAGAGCCATCAGAGACTCCAT 2610
566 tAsnLysLeuArgLysLysLeuGluArgArgArgGluArgGluAlaAla 583
2611 GAGCAACCTTAGAGAAAGGTTAGAGAGCGCTCGAAGGGAAGAGAGCTG 2660
583 spGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThrThr 599
2661 ACCAGGGTGGTTTGAAGGATGGTTCAACAGGTCCTCTGGATGACACC 2710
600 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuTh 616
2711 CTGCTTTCTGCTGACGGGACCCCTAGTAGTCTGCTCTCTGTACTTAC 2760
616 rValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgV 633
2761 AGTTGGGCTTGCCTTAATTAATAGGTTTGTGCTTTGTTAGAGACGAG 2810
633 alSerAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeuLeu 649
2811 TGAGTGAGTCCAGATCATGTACTTAGCAACAGTACCAAGGCGCTTCTG 2860
650 SerGlnGlyGluThrAspLeu 656
2861 AGCCAAGGAGAACTGACCTC 2881
seq_name: gb_v1:PERNV2
seq_documentation_block: 3482 bp RNA VRL 14-OCT-1997
LOCUS PERNV2 Porcine endogenous retrovirus env gene, 3482 bp.
DEFINITION Y12239 Porcine endogenous retrovirus
ACCESSION Y12239
VERSION Y12239.1 GI:2576324
KEYWORDS env gene.
SOURCE porcine endogenous retrovirus.
ORGANISM porcine endogenous retrovirus
REFERENCE 1 (bases 1 to 3482)
AUTHORS Le Vissier,P., Stoye,J.P., Takeuchi,Y., Patience,C. and Weiss,R.A.
TITLE Two sets of human-tropic pig retrovirus
JOURNAL Nature 389 (6652), 681-682 (1997)
MEDLINE 97478526
REFERENCE 2 (bases 1 to 3482)
AUTHORS Stoye,J.P.
TITLE Direct Submission
JOURNAL Submitted (02-APR-1997) J.P. Stoye, National Institute for Medical
REMARK Research, Virology, The Ridgeway Mill Hill, London, NW7 1AA, UK
FEATURES
source
1..3482
/organism="porcine endogenous retrovirus"
/specific_host="Sus scrofa"
/db_xref="taxon:61673"
/note="PERV-B"

```

gene

911.. .2884

/gene="env"

CDS

911.. .2884

/gene="env"

/codon_start=1

/protein_id="CAA72928.1"

/db_xref="GI:2576325"

/db_xref="SPTREMBL:O41173"

/translation="MHPTLSRHLPTRGGEKRPRLRPLSFASIAWFLTLTITPOASSK
RLDSSNHRPLSLTWLIIDPDGTVTNSTRGVAPRGTWPELHFCRLRLINPAVKSTP
PNLVRSGYVCCPGTEKEKYCGSGESFCRRSCVTSNDGDKWPISLQDRVKFSFN
SGPGLKMKLYDKSCSFSDVDLYLKISFETGQENIQKWINGMSWGIIVFYKGGGA
GSTLTIRLETGTEPPVAVGDPKVLAEQGPALPEPHNLVPQLTSLRPDIITQPSSN
GTGLIPTNPRNSPGVPVKTGQRLFSLQGAFAINSTDPAATSCWLCCLSSGPPY
EGMAKGGKFNVTKEHRNQCTWGRNKLTITLTVSGKGTCTGKAPPSHOHLCTVYVEQ
ASNYLYPCYNWAKNTGLTPCVSTSVFNOSKDFCVWQIVPRVYHPEEVLDEY
DYRNPKEPVSLLTAVMLGLTAVGVTGTAALITGPQOLEKGLGELHAAWTEDEL
ALEESVSNLEESTLSLEVYLNRRGLDLLFLREGGLCAALKEECFFYVDHSGAIRDS
MSKLRELRERRRREADQCFEGWFRNSPFWMTLLSALTGPLVLVULLLTIVGPCLIN
RFVAFVRERSAVQIMVLRRQOGLLSQGETDI."

BASE COUNT 927 a 854 c 867 g 834 t

ORIGIN

alignment_scores:

Quality: 3433.00 Length: 658
Ratio: 5.257 Gaps: 2
Percent Similarity: 99.240 Percent Identity: 98.176

alignment_block:

US-09-171-553B-6 x PERENV2 ..

Align seg 1/1 to: PERENV2 from: 1 to: 3482

1 MetHisProThrLeuSerArgArgHisLeuProThrArgGlyGlyGluPr 17
|||||
911 ATGCATCCACGTTAAGTCGGCGCCACCTCCCGACTCGGGGTGGAGACC 960
17 olysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuT 34
|||||
961 GAAAAGACTGAGAACCCCTTAAGCTTCGCCCTCCATCCCGTTCCTTA 1010
34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50
|||||
1011 CTCTAACATAACTCCCGAGCCAGCTAGCTAAACGCCCTTATAGACAGCTCG 1060
51 AsnProHisArgProLeuSerLeuThrTrpLeuIleIleAspProAspTh 67
|||||
1061 AACCCTCCATAGACCTTTATCCCTTACCTGGCTGATTTATGACCCTGATAC 1110
67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrpT 84
|||||
1111 GGGTGTCACTGTAATACACTCGAGGTGTGCTCCTAGAGGCACCTGGT 1160
84 rpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
|||||
1161 GGCTGAACCTGCAATTCGCTCCGATTGATTAAACCCCGCTGTAAAGAGC 1210
101 ThrProProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117
|||||
1211 ACACCTCCCAACTAGTCGTAGTTATGGGTTCTATTGCTGCCAGGCAC 1260
117 rGluLysGluLysTyrCysGlySerGlyGluSerPheCysArgArgT 134
|||||
1261 AGAGAAGAGAATAACTGTGGGGTCTGGGAATCCCTTCGTAGGAGAT 1310
134 rpSerCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150
|||||
1311 GGAGCTGGCTCACCTCCCAACGATGAGACTGGAAATGCCGATCTCTCTC 1360
151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrLy 167
|||||
1361 CAGGACCGGGTAAAAATTCCTTTTCTCAATTCCGCGCCGGGCAAGTACAA 1410

167 sMetMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAspT 184
|||||
1411 AGTGATGAACATATATAAGATAAGAGCTGCTCCCATCAGACTTAGATT 1460
184 yrLeuLysIleSerPheThrGlu..ArgLysThrGlyLysTyrSerLysV 200
|||||
1461 ATCTAAAGATAAGTTTCACTGAAAGAGAAACAGGAAATATATCAAAAG 1510
200 alaAspLysTrpTyrGluLeuGlyAsnSerPheLeu..LeuTyrGlyGly 216
|||||
1511 TGGATAAATGGTATGAGCTGGGAATAGTTTTTATAAATATGCGGGGG 1560
216 yAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGluP 233
|||||
1561 AGCAGGCTCCACTTTAAACCATTCGCCTTAGAGTAGAGACGGGACAGAAC 1610
233 roProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProPro 249
|||||
1611 CCCTGTGGCAGTGGGACCGCATAAAGTACTGGCTGAACAGGGGCCCCC 1660
250 AlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuAr 266
|||||
1661 GCCCTGGAGCCCGCATAACTTGGCGGTGCCCAATTAACCTCGCTGCG 1710
266 gProAspIleThrGlnProProSerAsnSerThrThrGlyLeuIleProT 283
|||||
1711 GCCTGACATAACACAGCGCGCTAGCAACGGTACCCTGGATTGATTCTTA 1760
283 hrAsnThrProArgAsnSerProGlyValProValLysThrGlyGlnArg 299
|||||
1761 CCAACAGCGCTAGAAACTCCCGAGGTTCCTGTCTTAAGACAGGACAGAG 1810
300 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspPr 316
|||||
1811 CTCCTTCAGTCTATCCAGGAGGCTTTCCAAAGCCATCAACTCCACCGACC 1860
316 oAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTyrT 333
|||||
1861 TGATGCCACTTCTTCTTGTGGCTTCTCTATCTCCTCAGGGCCCTCTATT 1910
333 yrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArg 349
|||||
1911 ATGAGGGGATGGCTAAAGAGGAAATTCATGTGACCAAGAGACATAGA 1960
350 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSe 366
|||||
1961 AATCAATGTATAGTGGGGTCCCGAAATAAGCTTACCCTCCTCAAGTTTC 2010
366 rGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeuC 383
|||||
2011 CGGGAAGGGACATGCATAGSAAAGCTCCCCCATCCCAACACCTTT 2060
383 ystYrSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeuVal 399
|||||
2061 GCTATAGTACTGTGTTTATGAGCAGGCCTCAGAAAAATCAGTATTAGTA 2110
400 ProGlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysVa 416
|||||
2111 CCTGGTTATACAGGTGGTGGGATCAATCACTGGGTAAACCCCTGTGT 2160
416 lSerThrSerValPheAsnGlnSerLysAspPheCysValMetValGlnI 433
|||||
2161 TTCACCTCAGCTCTCAACCAATCCAAAGATTCTGTGTCTATGGTCCAAA 2210
433 leValProArgValTyrTyrHisProGluGluValValLeuAspGluTyr 449
|||||
2211 TCGTCCCGGAGTGTACTACCATCTCCTGAGGAAGTGGTCTTGATGAATAT 2260
450 AspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAl 466
|||||
2261 GACTATCGGTATATACCGACCAAAAGAGAACCCCGTATCCCTTACCCTAGC 2310

466 aValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAlaA 483
 2311 TGTAAATGCTCGGATTAGGACGGCGTTGGCGTAGGAACAGGACAGCTG 2360
 483 laLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuGlyGluLeuHis 499
 2361 CCTGATCACAGGACACACAGCTAGAGAAAGGACTTGGTAGCTACAT 2410
 500 AlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnLe 516
 2411 GCGGCCATCACAGAAGATCTCCGAGCCTTAGAGGAGTCTGTTAGCAACCT 2460
 516 uGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgG 533
 2461 AGAAGAGTCCCTGACTCTTTGTCGAAGTGGTCTACAGAACCGAGGG 2510
 533 lyLeuAspLeuLeuPheLeuArgGlyGlyLeuCysAlaAlaLeuLys 549
 2511 GATTAGATCTGCTTTCTAAGAGAGGTGGTTATGTGCAGCCTTAAAA 2560
 550 GluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMe 566
 2561 GAAGAATGTTGCTTCTATGTAGATCACTCAGGAGCCATCAGAGACTCCAT 2610
 566 tAsnLysLeuArgLysLysLeuGluArgArgArgGluArgGluAlaA 583
 2611 GAGCAAGCTTACAGAAAGGTTAGAGAGCGTCGAAGGAAAGAGAGGCTG 2660
 583 spGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThrThr 599
 2661 ACCAGGGTGGTTTCAAGGATGTTCAACAGCTCTCCTTGGATGACCACC 2710
 600 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeuTh 616
 2711 CTGCTTCTGCTCTGACGGGACCCCTAGTAGTCTGCTCTGCTTACTTAC 2760
 616 rValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgV 633
 2761 AGTGGGCTCTCTTAATAATAGTTTGTGGCTTTGTTAGAGAACAGAG 2810
 633 alSerAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeuLeu 649
 2811 TCAGTGCAGTCAGATCATGTTAGTACCTAGGCAACAGTACCAAGGCTCTG 2860
 650 SerGlnGlyThrAspLeu 656
 2861 AGCCAAGGAGAACTGACCTC 2881

seq_name: gb_vi: PEN293657

seq_documentation_block: 30-MAY-2001
 LOCUS PEN293657 8763 bp DNA VRL
 DEFINITION Porcine endogenous type C retrovirus proviral gag gene, pol gene and env gene, class B, clone 213.
 ACCESSION AJ293657
 VERSION AJ293657.1 GI:14275841
 KEYWORDS env gene; envelope; gag gene; group specific antigen; pol gene; polymerase; polyprotein.
 SOURCE porcine endogenous type C retrovirus.
 ORGANISM porcine endogenous type C retrovirus.
 Viruses: Retroid viruses; Retroviridae; Gammaretrovirus.
 REFERENCE 1 (bases 1 to 8763)
 AUTHORS Krach, U., Fischer, N., Czauderna, F. and Tonjes, R.R.
 TITLE Comparison of replication-competent molecular clones of porcine endogenous retrovirus class a and class b derived from pig and human cells
 JOURNAL J. Virol. 75 (12), 5465-5472 (2001)
 MEDLINE 21256017
 REFERENCE 2 (bases 1 to 8763)
 AUTHORS Tonjes, R.R.
 TITLE Direct Submission
 JOURNAL Submitted (31-JUL-2000) Tonjes R.R., Medical Biotechnology, Paul-Ehrlich-Institut, Paul-Ehrlich-Str. 51-95, Langen, 63225,

GERMANY

FEATURES
 source
 Location/Qualifiers
 1..8763
 /organism="porcine endogenous type C retrovirus"
 /proviral
 /specific_host="Sus scrofa"
 /db_xref="taxon:70540"
 /clone="213"
 /cell_line="PK15"
 /note="class B"
 1..630
 467
 /note="Cap site"
 633..650
 1077..2651
 /gene="gag"
 1077..2651
 /gene="gag"
 /codon_start=1
 /product="group specific antigen"
 /protein_id="CAC39625.1"
 /db_xref="GI:14275842"
 /translation="MGQTVTTPLSLTLDHMTVRSRAHNSLVQVKKGPWQTFCASEWP
 TFDGWPSEGTNSFELILAVKAIIFOTGPGSHPDQEPVILTWDLAEPWPVVKPWLN
 KPRKGPRIALGKKNKHSAEKVPSEGLIYEIPEPTWPEPQVPPPPYPAQGNVC
 PSAPPKVPVEGPAAGTSRRGATPERIDEAILPRITGYPFGGQLOPQIPFFSS
 ADLYNMKTHNPPFSEDQRLTGLVESLMSHOPTWDDCQQLQTLFTFEERILILEA
 RKNVPGADRGPTQLQNEIDMGFPLTRPGWDYNTAEGRESLKIYQALVAGLSRRP
 TNLAKREVMPGPNPEPSVFLERLMEAFRRFTFPDPTSEAKSVALAFIGQSALDIR
 KKLORLEGLQAEULDLVREAEKVYYRRETEKEPQKEKEREREERDRRQKNLT
 KILAAVEGKSRSRERDFKIRSGPRQSGNLNRTPLDKDQCAYPEKKGHWARNCPK
 KNGKPKVLALEEDKD"
 2652..6239
 /gene="pol"
 <2652..6239
 /gene="pol"
 /note="complete pol ORF, gag-pol is expressed as a polyprotein"
 /citation=[1]
 /codon_start=1
 /product="polymerase"
 /protein_id="CAC39626.1"
 /db_xref="GI:14275843"
 /translation="GRKSDPLPEPRVTLKVEGQPVFELVDTGAHSHVLLQPLKKE
 KKSWMGATGQRYQPTTRIVDLAVGRVTHSFVIPCVPPLIGRDLITKMGQISF
 EQGREVSNKPIITVLTLDDEYRLYSQVKPDQDIQSWLEQFPQQAETAGMGLA
 KQVPQVILKASATPVSVQYPLSREAREGIMPHVQRLIQOGLVQVSPWNTPLLP
 VRKGTNDYRPVQADLRVNRKVDIHTVNPYNLLSALPERNWYVLDLDAFFCL
 RLHPTSOPLAFENRDPCTGRTGOLTWRLPQGFKNSTPFDEALHDLANFRHQPC
 VTLQYVDDLLAGATKQDCLEGTALLLESLDGLYRASAKAQICREVTYLYSLR
 GQRWLTEARRKVVQIPAPITAKQVREFTAGFCRWIPGFATLAAPLPLTKKRG
 EFSWAPHOXKAPDAIKKALLSAPALPDVTKPTLYYDERKGVARGVLTQTLPWRR
 PVAYLSKLDPVAGWPVCLKAIAVALIVKADKLTIGQNTVIAHALNIVRQPP
 DRWNTNARWTHYQSLLTPEVTFAPPAALNPATLLPEETDEPTDHCHQLIEETGVR
 KDLDITPLTGEVLTWFTDGSYVVEGKMAAAVVDGTRTWASSLPGHTSAQAEML
 ALTOALRLAEGKSNINYDYSKAFATAHVHGAIVKQGLLTSAERLKNKEILSLE
 ALHLPKRLAIIHCPHQRAKDLISRGNMADRVAKQAQAAVNPILLTIETPKAPRRQ
 YTLDEWQEKIKDQFSEPTGCTSYGKEITLPHKLEGLEYVQQLHRLTHLKHQQL
 VRTSPYHVLRLPGVADSVKVCQVNPANPSRIPGKRLGRSGHPGAHWEDFTEVK
 PAKYGNKLVLEVDTFGWVEAYPTKSTVAVAKILEEIFPRFGIPKVGSDNGPA
 FVAOVSGLAKILGIDMKLHCAYRPOSSGOVVRNMRTIKETLTKLTTGTINDWIAL
 PFVLPFRVNTPGQGLTPYELLYGPPPLVEIASVHADVLISQPLSRKALEWVRQ
 RANKOLREAYSGEGDLQVPHRFQVGDSDSVYRRHHAGNLETRWKGPLYLVLLTPTAVKY
 EGISWIHASHVKKLAPPDSDGWRAEKTEKNLKLRLHLRVPSYNNNSPQ"
 6112..8085
 /gene="env"
 6112..8085
 /gene="env"
 /codon_start=1
 /product="envelope"
 /protein_id="CAC39627.1"
 /db_xref="GI:14275844"
 /translation="MHPTLSWRHLPTRGGEPKRLRIPLSFASIAWFLTLTITPQASSK

RLIDSSNPHRPLSLTLIIDPDGTGVNSTRGVAPRGTWPELHFCRLINPAVKSTP
 PNLVRSYFCYCCPTEKEKYGSGSESCREWSQVTSNDGDKWPIQLQDRVKFSYFN
 SGPFIKVMKLIKDKSCSPSBLDYLKISFTEKGRQENIKQWINGMSWGIYVYKYGGA
 GSTLIRLRIETGTPEPPVAVGPKVLAEGPPALEPPHNLPEVQLTSLURPDITPPSX
 STTGLIPTNTPRNSPVGKTKQRLFSLIQAFQAINSTDPAATSSWLCILSSGPPYY
 ECKMAKGFNVTKEHNRQCTWGSRNKILTLTEVSGKTCIGKAPPSHQLCNSTVYVEQ
 ASENOLYVPGYNRMWACNTGLTPCVSTSVFNOSKDFCVMOIIVPRVYHPEEVLDEY
 DYRYNRKPREPVSLTLAVMLGLTAVGTCTAALITGTPQOLEKGLGELHAAETDLR
 ALLESVSNLESLTSLSEVLIQNRGLDLLPLREGGGLCAALKKECCFYVDHSGAIRS
 MSKLRLERLRRRREADQGFEGFNKSPNMTLLLSALTGLPLVVLLLLLLLVGVCLIN
 RFVAFVRERVSVAQIVMLRQYQGLLSQGETDL*

17r polyA_signal 8133..8763
 BASE COUNT 2301 a 2200 c 2287 g 1974 t 1 others
 ORIGIN

alignment_scores:
 Quality: 3421.00 Length: 658
 Ratio: 5.255 Gaps: 2
 Percent Similarity: 98.936 Percent Identity: 98.024

alignment_block:

US-09-171-553B-6 x PEN293657 ..

Align seg 1/1 to: PEN293657 from: 1 to: 8763

1 MethisProThrLeuSerArgArgAlaHisLeuProThrArgGlyGlyGluPr 17
 6112 ATGCATCCCAAGCTTAAGTCGGCCACCTCCCGACTCGGGGTGGAGACC 6161

17 oLysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuT 34
 6162 GAAAGACTGGAATCCCTTAAGCTTCGCCCTCCATCGCCUGGTCCCTTA 6211

34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50
 6212 CTCTAACAAATAACTCCCGAGCCAGTAGTAAACGCCCTATAGACAGCTCG 6261

51 AsnProHisArgProLeuSerLeuThrTrpLeuIleIleAspProAspTh 67
 6262 AACCCCATAGACCTTTATCCCTTACCTGGCTGATTATTGACCCCTGATAC 6311

67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrpT 84
 6312 GGGTGTCTACTTAATAGCACTCAGGTGTGCTCTTAGAGCACTGGT 6361

84 rpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
 6362 GGCCTGAACCTGATTTCTGCTCCGATTGATTAAACCCCGCTGTTAAAGC 6411

101 ThrProProAsnLeuValArgSerThrGlyPheThrCysCysProGlyTh 117
 6412 ACACCTCCCAACCTAGTCCGAGTATAGGGTCTATTGCTCCCGAGGCAC 6461

117 rGluLysGluLysThrCysGlyGlySerGlyGluSerPheCysArgArgT 134
 6462 AGAGAAAGAGAAATACTGTGGGGTCTGGGGAATCTCTCTGTAGGAGAT 6511

134 rpSerCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150
 6512 GGAGCTGCGTCACCTCCCAACGATGGAGACTGGAAATGGCCGATCTCTCTC 6561

151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysThrLy 167
 6562 CAGGACCGGGTAAATCTCTTTGTCAATTCGCGCCCGGCAAGTACAA 6611

167 sMetMetLysLeuThrLysAspLysSerCysSerProSerAspLeuAspT 184
 6612 AGTGATGAACATATATAAGATAAGAGCTGCTCCCATCAGACTTAGATT 6661

184 yrLeuLysIleSerPheThrGlu..ArgLysThrGlyLysThrSerLysV 200
 6661

6662 ATCTAAAGATAAGTTTCACTGAAAAAGGAAAAACAGAAATATTCAAAG 6711

200 alAspLysTrpThrGluLeuGlyAsnSerPheLeu..LeuThrGlyGlyG 216
 6712 TGGATAAATCGGTATAGCTGGGGAATAGTTTTTTTAAATATGCGGGGG 6761

216 yAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGluP 233
 6762 AGCAGSGTCCACTTTAAACCATTCGCTTAGGATAGAGACGGGACAGAAC 6811

233 roProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProPro 249
 6812 CCCCTGTGCAGTGGACCCGATAAAGTACTGGCTGAACAGGGGCCCCGG 6861

250 AlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuAr 266
 6862 GCCTGGAGCCACCGCATAACTTGGCGGTGCCCAATTAACCTCGCTGGG 6911

266 gProAspIleThrGlnProProSerAsnSerThrThrGlyLeuIleProT 283
 6912 GCCTGACATAACACAGCGCTTAGCANCACTACCACTGGATTGATTCCTA 6961

283 hrAsnThrProArgAsnSerProGlyValProValLysThrGlyGlnArg 299
 6962 CCAACACGCTAGAACTCCCGAGGTGTTCTCTTAAAGACAGACAGAGA 7011

300 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspPr 316
 7012 CTCTTCAGTCTCATCCAGGAGCTTCCAGGCCATCACTCCACCGACCC 7061

316 oAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTyrt 333
 7062 TGATGCCACTTCTTCTTGTGGCTTGTCTCTCCTCAGGCGCTCCTTAT 7111

333 yrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArg 349
 7112 ATGAGGGGATGGCTAAGAGGAAATCAATGTGACCAAGACAGATAGA 7161

350 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuGluValSe 366
 7162 AATCAATGTACATGGGGTCCCGAATAAGCTTACCTCCTCAGTGAAGTTTC 7211

366 rGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeuc 383
 7212 CGGGAAGGGGACATGCATAGGAAAGCTCCCGCATCCCAACACCTTT 7261

383 ystyrSerThrValValThrGluGlnAlaSerGluAsnGlnThrLeuVal 399
 7262 GCAATGACTGTGGTTTATGACGAGGCTCAGAAATCAGTATTAGTA 7311

400 ProGlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysVa 416
 7312 CCTGGTTAATACAGGTGGTGGCATCAATCTGCTTAACCCCTGTGT 7361

416 lserThrSerValPheAsnGlnSerLysaspPheCysValMetValGlnI 433
 7362 TTCACCTCATCTCTTCAACCAATCCAAAGATTCTGTGTCATGGTCCAAA 7411

433 leValProArgValThrHisProGluGluValValLeuAspGluTyr 449
 7412 TCGTCCCCCGAGTGTACTACCATCCCTGAGGAAGTGGTCTGTGATGATAT 7461

450 AspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAl 466
 7462 GACTATCGGTATAACCGACCAAAAGAGAACCCGCTATCCCTTACCCCTAGC 7511

466 aValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAlaA 483
 7512 TGTAAATGCTCGATTAGGGACCGCGTGGCTAGGAACAGGACAGCTG 7561

483 laLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuGlyGluLeuHis 499
 7562 CCCTGATCACAGGACACACAGCTAGAGAAAGGACTTGGTGAGCTACAT 7611

500 AlaAlaMetThrGluAspLeuArgAlaLeuLeuLysGluSerValSerAsnLe 516
|||||
7612 GCGCCATGACAGAAAGATCCGAGGCTTAGAGGAGTCTGTAGCAACCT 7661
516 uGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgG 533
|||||
7662 AGAAGAGTCCCTGACCTCTTGTCTGAAGTGGTCTACAGAACCGGAGG 7711
533 lYLeuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaLeuLys 549
|||||
7712 GATTAGATCTGCTGTTTCTAAGAGAAGTGGGTATGTGAGCCTTAAAA 7761
550 GluGluCysCysPheTrValaspHisSerGlyAlaIleArgAspSerMe 566
|||||
7762 GAGATGTTGCTCTATGTAGATCACTCAGGAGCCATCAGAGCTCCAT 7811
566 tAsnLysLeuArgLysLeuGluArgArgArgGluArgGluGluAla 583
|||||
7812 GAGCAAGCTTAGAGAAAGGTTAGAGAGCGCTCGAAGGGAAGAGAGGCTG 7861
583 spGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThrThr 599
|||||
7862 ACCAGGGGTGGTGTGAAGGATGTTCAACAGGTCTCTGGATGACCACC 7911
600 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuTh 616
|||||
7912 CTGCTTCTGCTGACGGGGCCCCCTAGTAGTCTCTGCTCTTACTTAC 7961
616 rValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgV 633
|||||
7962 AGTTGGGCTTGTGTAATTAATAGTGTCTGCTCTTGTAGAGAACGAG 8011
633 aISerAlaValGlnIleMetValLeuArgGlnGlnTrpGlnGlyLeuLeu 649
|||||
8012 TGAGTGCACTGACATCATGTTAGTACTTAGGCAACAGTACCAAGGCTTCTG 8061
650 SerGlnGlyGluThrAspLeu 656
|||||
8062 AGCCAGGAGAACTGACCTC 8082

seq_name: gb_v1:PERY17013

seq_documentation_block:
LOCUS PERY17013 7808 bp DNA VRL 14-APR-2000
DEFINITION Porcine endogenous retrovirus gag, pol, env genes.
ACCESSION Y17013
VERSION Y17013.1 GI:3005985
KEYWORDS env gene; gag gene; pol gene.
SOURCE porcine endogenous retrovirus.
ORGANISM porcine endogenous retrovirus
Viruses; Retroid viruses; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
1 (bases 1 to 7808)
Czauderna,F., Fischer,N., Boller,K., Kurth,R. and Tonjes,R.R.
Establishment and characterization of molecular clones of porcine
endogenous retroviruses replicating on human cells
J. Virol. 74 (9), 4028-4038 (2000)
20219394
2 (bases 1 to 7808)
Toenjes,R.R.
Direct Submission
Submitted (30-MAR-1998) R.R. Toenjes, Paul-Ehrlich Institut,
Paul-Ehrlich Str. 51-59, P.O.Box 1740, D-63207 Langen, FRG

FEATURES

source
1. .7808
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
/tissue.type="kidney"
/cell_line="PK15"
260. .1834
/gene="gag"
260. .1834

gene
CDS

/gene="gag"
/codon_start=1
/protein_id="CAA76581.1"
/db_xref="GI:3005986"
/db_xref="SPTREMBL:O73504"
/translation="MGQTITPLSLTLDHWTVEVRSAHLNLSVQVKGKPMQTPCSEWEP
TFDWPSEGTENSEILAVAKIIFQTSPSHPDQEPYILTWDLAEDPPVVKWKLNL
KPKPGPRILALGKHNKSAKEVSPRIYETIEPTWPEQPPVPAQVAVRG
PSADPGAVVEGPAAGTRSRGATPERTDEAILPLRTYPPMPGQLOPLQWPFSS
ADLYNMKTWNHPPESDPORLTGLVSLMFSHQPTWDDCOOLLOTLFTAEERILLEA
RKNPCADGRPTOLONEIDMGFPLRLWDYNTAEGRESLKIRYALVAGLRGASRRP
TNLAKRVNQGFNEPPSVFLERLMEAFRRFTFPDPTSEAKASVALAFIGOSALDIR
KKLQRLGLEQAEELRDLVREAKEYYRETEEEKERKEEREEREERDRRQEKNT
KILAAVVEGKSSRRERDRFRKIRSGPROSGNLNRTPLDKQDCAYCKCKGHWARNCPK
2120. .4726
/gene="pol"
2120. .4726
/gene="pol"
/codon_start=1
/protein_id="CAA76582.1"
/db_xref="GI:3005987"
/db_xref="SPTREMBL:O73505"
/translation="MGAQISFEQGREPEVSNKKNPITVITLQLDDERYLSQVKKPDQD
IOSWLEFPQAWAETAGMLAKOPQVILKASATPSVSRQYPLSRAREGINPHVQ
RLIQGILVYQSPWNTPLLPVRKEGTINDYRVQDLREVNRKRVQDIHPTVNPYLLS
ALPPERNMTYVLDLDAFFCLRLHPTSQPLFAFWRDPDGTGRTGTLRLPQGFKN
PTIFDEALHRDLANFRIOHPQVTLQYVDDLLLAGATKQDCLEGTKALLLESLDLYR
ASAKAQCRRREVTVLGYSLRGGRHLTEARKTVVQIAPATKQVREFLGTAGFCR
LWIPGFATLAAPLYPLTKEGFEFSAPEHQAFDAIKALLSAPALAPDVTKPTLY
VDERGTVAGVLTQTLGPWRPVAYLRSKLDPAVSGWPVCLKAIAAVALIKVDADKLT
LGONTIAPHALENIVRQPPDRWMTNARMTHYQSLLLETRVTFAPALAPATLPE
ETDEPVTDCQHLLEETGVRKDLTDIPLTGEVLTWFTDGSYVVEGKRMAGAAVVDG
THIWASSLPCTSAQKALMALTOALRLABKSINIYTDYAFATAHVGAIVKOR
GLTSAGREIKNKERILSLLEALHLPKRLIITHCPGOKAKDLISRGQMDRVAKQA
AGAVNLPIETPRAPRPROYTLLEDQEIKKIDOFSETPGTCYTSYKGEILLPHKEG
LEYVQIIRHLHLGTHLQQLVRSFYHVRDGPVADSVVAHCVQCLVNANPSKIPP
GKRLGSHPGAHWEVDFTEVVPKYGKYLVLVDFTSFGWVEAYPTKKTSTVVAKKY
WRKFFRDLEYLR"
5294. .7267
/gene="env"
5294. .7267
/gene="env"
/codon_start=1
/protein_id="CAA76583.1"
/db_xref="GI:3005988"
/db_xref="SPTREMBL:O73506"
/translation="MHPITLSWRHLPTRGGEPRKLRIPLSFASIANFLTTLTPQASSK
RLIDSSNHRPLSLTLWLIIDPDTGYTNSTRGVAPRGTWPFELFCLRLNPAKSTP
PNLRSYGYCCPCKEKYCGGSEESFCRWSCVTNDGDKWPIQLQDRKFSFVN
SGPKYKVMKLYKDKSCPSDLYLKISFTEKQENIKQWINGMSWGVIFYKYGGGA
GSTLTIRLRIETGTEPPVAVGPKVLAQGGPAPLEPPHNPPLVPOLTLRPDITQPSN
GTTGLIPTNTPNSGVPVKTGORLFSLIQGAFOAINSTDPDATSSCWLCLSGGPYV
EGMAKEGFNVTKEHKNCTWGSBNKLTLEVSCKGTCIGKAPSHOHLCLYSTVYEQ
ASENQYLPVGNRWACNTGLTPCVSSSVFNQSCDFVMVOILPRVYTHPEVVIDEY
DYRNRPAREPVSLLTAVMLGLGTAVGTGTAALITGPOLERKGLGELHAAMTEDLR
ALEPSVNLDESLSLSEVLQNRGLDLEFREGGLCAALKKECCYFVDRHSGAIRS
MSKLRLERRRRERDOGWFEGRFNRSPMNTLLSALTGLPLVLLLLLTGVCCLIN
RFVAFVRVRGAVOIMVLROOYQGLLSQGETDL"
BASE COUNT 2077 a 1971 c 2047 g 1713 t
ORIGIN

alignment_scores:

Quality: 3417.00 Length: 658
Ratio: 5.241 Gaps: 2
Percent Similarity: 99.088 Percent Identity: 97.720

alignment_block:

US-09-171-553B-6 x PERY17013

Align seg 1/1 to: PERY17013 from: 1 to: 7808

```

1 MethSProThrLeuSerArgArgHisLeuProThrArgGlyGlyGlnP 17
|||||
5294 ATGCATCCACGTTAAGCTGGGCGACCTCCCGACCTCGGGGTGGAGAGCC 5343
17 GlysArgLeuArgIleProLeuSerPheAlaSerIleAlaIrrPheLeu 34
|||||
5344 GAAAGACTGGAATCCCTTAAGCTTCCCTCCATCCGCTGGTTCCTTA 5393
34 hrLeuThrIleThrProGlnAlaSerSerIleAlaIrrPheLeu 50
|||||
5394 CTCTACACATATACTCCCGAGCGCACTAGTAACCGCTTATAGACAGCTCG 5443
51 AsnProHisArgProLeuSerLeuThrIrrPheLeuIleAlaProAspTh 67
|||||
5444 AACCCCATAGACCTTTATCCCTTACCTGCGCTGATATTAGACCTGATAC 5493
67 rGlyAlaIrrValAsnSerThrArgGlyValAlaIrrArgGlyIrrThrP 84
|||||
5494 GGGTGTCACTGTAATAGCACTCGAGGTGTTGCTCCAGAGCACTGGT 5543
84 rPProGlnLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
|||||
5544 GGCCGAGACTGATTTCTGCTCCGATGATTAAACCGCTGTTAAAAAGC 5593
101 ThrProProAsnLeuValArgSerTyrGlyPheTyrCysProGlyTh 117
|||||
5594 ACACCTCCCAACCTAGTCCGTAGTTATGGGTCTATTGCTGCCAGGAC 5643
117 rGlyArgGlyLysTyrCysGlyGlySerGlyGlySerPheCysArgArgT 134
|||||
5644 AGAGAAAGAAATACCTGTGGGGTCTGAGGAATCCCTGTGTGAGAGAT 5693
134 rPSerCysValThrSerAsnAspGlyAspTrrPlyStrProIleSerLeu 150
|||||
5694 GGAGCTGGCGTACCTCCACAGATGGAGCTGGAATGGCCGATCTCTCTC 5743
151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrLy 167
|||||
5744 CAGGACCGGGTAAATCTCTTGTCAATCCGCGCCGCGCAAGTACAA 5793
167 SmetMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuSpt 184
|||||
5794 AGTGATGAACCTATATAGATTAAGAGCTGCTCCCAATGACACTTAATT 5843
184 TyrLeuLysIleSerPheThrGln..ArgLysThrGlyLysTyrSerLysV 200
|||||
5844 ATCTAAGATTAAGTTTCTACTGAAAGAAAGAAACAGGAATATTCANAAG 5893
200 AlaAspLysTrrPyrGlyLeuGlyAsnSerPheLeu..LeuTyrGlyGly 216
|||||
5894 TGGATTAATGGTATGAGCTGGGGAATAGTTTATTAATATATGCGGGG 5943
216 ValGlySerThrLeuThrIleArgLeuArgIleGlyThrGlyThrGlnP 233
|||||
5944 AGCAGGCTCCACTTAAACATTCGCCCTTAGGATAGAGACGGGACAGAAC 5993
233 rOProValAlaMetGlyProAspLysValLeuAlaGlnGlnGlyProPro 249
|||||
5994 CCCCTGGGAGTGGGAGACCGCATTAAGTACGGCTGAACAGGGGCCCGCC 6043
250 AlaLeuGlnPProProHisAsnLeuProValIProGlnLeuThrSerLeuAr 266
|||||
6044 GCCCTGGAGCCACCGCATTAATCTCCGCTGCCCAATTAACCTCGCGCG 6093
266 gProAspIleThrGlnPProProSerAsnSerThrThrGlyLeuIleProT 283
|||||
6094 GCCTGACATATACACAGCCGCTTACCAACGATACCATGGAATGATTCCTA 6143
283 hrAsnThrProArgAsnSerProGlyValIProValLysThrGlyGlnArg 299
|||||
6144 CCAACAGCGCTTGAAGAACTCCCGAGGTGTCCTGTTAAGACAGACAGACAG 6193
300 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspTr 316
|||||
|||||
6194 CTCTTCAGTCTCATCCAGGAGGCTTTCAGACCATCAACTCCACCGACC 6243
316 oAspAlaIrrThrSerSerCysTrrPleuCysLeuSerSerGlyProProTyr 333
|||||
6244 TGATGCCACTCTCTTGTGGCTTGTGTATCTTCACAGGGCTTCCTTATT 6293
333 YrGlnGlyMetAlaIyGlnArgLysPheAsnValThrLysGlnHisArg 349
|||||
6294 ATGAGGGGATGGCTAAAGAAAGAAATTCATATGTGACCAAGAGCATAGA 6343
350 AsnGlnCysThrTrrPglySerArgAsnLysLeuThrLeuThrGlyValSe 366
|||||
6344 AATCAATGTACATGGGGGTCCCGAATATAGCTTACCTCATCGAAGTTTC 6393
366 rGlyLysGlyThrCysIleGlyLysAlaIrrProProSerHisGlnHisLeu 383
|||||
6394 CGGGAAGGGGACATGCATAGGAAAGCTCCCATCCACCAACACCTTT 6443
383 YstrSerThrValIrrValIrrGlnAlaSerGlnAsnGlnTrrLeuVal 399
|||||
6444 GCTTATACTGTGTGTTATGACGAGGCTCAGAAAATCAGTATTAGTA 6493
400 ProGlyTrrAsnArgTrrPrrAlaCysAsnThrGlyLeuThrProCysVa 416
|||||
6494 CTGTGTTATACAGGTGGTGGCATGCAATACGTGGTTAACCCCTGTGT 6543
416 LserThrSerValPheAsnGlnSerLysAspPheCysValMetValGlnI 433
|||||
6544 TTCTCTCTCAGTCTTCAACCAATCCAAAGATTTCTGTGATAGTCCAA 6593
433 lValIrrProArgValIrrTrrHisProGlnGlnValValLeuAspGlyTyr 449
|||||
6594 TCGTCCCGCAGGTACTACTACCTCGAGAAATGCTCTTGTATGATAT 6643
450 AspTrrArgTrrAsnArgProLysArgGlnProValSerLeuThrLeuAl 466
|||||
6644 GACTATCGGTATTAACGACCAAAAGAAAGAACCCGTATCCCTTACCTAC 6693
466 aValMetLeuGlyLeuGlyThrIrrAlaValGlyValGlyThrGlyThAla 483
|||||
6694 TGTAAATGCTCGATTAGGAGCGCGCTTGGCTGAGGAAGAGGACGCTG 6743
483 lAlaLeuIleThrGlyProGlnGlnLeuGlnLysGlyLeuGlnHis 499
|||||
6744 CCTGTATCAGAGACCAACAGCACTTAGAAGAGACTTGGTAGCTACAT 6793
500 AlaAlaMetThrGlnAspLeuArgAlaLeuLysGlnSerValSerAsnLe 516
|||||
6794 GCGGCATGACAGAGATCTCCGAGCTTAAAGAGATGTTAGTACAACT 6843
516 uGlnGlnSerLeuThrSerLeuSerGlnValValLeuGlnAsnArgArg 533
|||||
6844 AGAAGAGTCCCTGACTTCTTGTCTGAAGTGTCTTACAGAACCGGAGGG 6893
533 lYLeuAspLeuLeuPheLeuArgGlnGlyLysCysAlaAlaLeuLys 549
|||||
6894 GATTATATCTGCTGTTTAAAGAGAGGTGGTTATGTCAACCTTAAAA 6943
550 GlnGlnCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMe 566
|||||
6944 GAAGAAATGTTGCTTATATAGATCACTCAGAGGACCATGAGACATCAT 6993
566 lAsnLysLeuArgLysLeuGlnGlnArgArgArgArgGlnAlaAla 583
|||||
6994 GAGCAAGCTTAGAGAAAGCTTAGAGAGGCTGAGAGGAAAGAGGCTG 7043
583 spGlnGlyTrrPheGlnGlyTrrPheAsnArgSerProTrrMetTrrThr 599
|||||
7044 ACAGAGGGGTGTTGAAGATGTTTCAACAGGTCTCTGATGACACAC 7093
600 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuThr 616
|||||

```



```

6249 AACCCCATGAGACCTTATCCCTTACCTGGCTGATTTATGACCCCTGATAC 6298
67 rglValAlaThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84
6299 GGGGTGACAGTAAATAGCACTGAGGTGTGCTCTAGAGCAGCACTGGT 6348
84 rpproGluLeuHisPheCysLeuArgLeuLeuLeuProAlaValLysSer 100
6349 GGCTGAACTGATTTCTGCTCCGATTTGATTACCCGCTGTAAAGC 6398
101 ThrProAsnLeuValArgSerTrpGlyPheTrpCysCysProGlyTh 117
6399 ACACCTCCCAACCTAGCTAGTATGGTTGTTCTATTGCTGCCAGCAGC 6448
117 rglLysGluLysTrpCysGlyLysSerGlyLysPheCysArgArgTr 134
6449 AGGAAAGAGAAATACTGTGGGGTTCTGAGGAATCCTCTGTAGAGAT 6498
134 rpsrCysValThrSerAsnAspGlyAspTrpLysTrpProLysSerLeu 150
6499 GGAGCTGCGTCACCTCCAAAGATGGAGACTGGAAATGGCCGATCTCTTC 6548
151 GluAspArgValLysPheSerPheValAsnSerGlyProGlyLysTrp 167
6549 CAGGACCGGGTAAATTTCTCTTGTCAATTCGGCCCGGCAAGTACAA 6598
167 smetMetLysLeuTrpLysAspLysSerCysSerProSerAspLeuAsp 184
6599 AGTGATGAACCTATATAAGATAGAGCTGCTCCCATCAGACTTAGATT 6648
184 rylLeuLysIleSerPheThrGluArgLysThrGlyLysTrpSerLys 200
6649 ATCTAAGATTAAGTTTCACTGAAAGGAAACAGAAATATTCAAAG 6698
200 AlaAspLysTrpTrpGlyLeuGlyAsnSerPheLeuLeuTrpGlyGly 216
6699 TGGATTAATGATAGCTGGGGGAATGTTTATTAATATGCGGGGCG 6748
216 yAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyTrpGlu 233
6749 AGAGGGTCCACTTAACCATTCGCTTAAGATAGAGAGGGGAGACAGAC 6798
233 rorProValAlaMetGlyProAspLysValLeuAlaGluGlnLysPro 249
6799 CCCCTGTGGCAGTGGACCCGATTAAGTACTGTGAACAGGGGCCCGCG 6848
250 AlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeu 266
6849 GCCCTGGAGCCACCGCATTAAGTGGCGGCCCAATTACCTCGCTCG 6898
266 gProAspIleThrGlnProProSerAsnSerThrThrGlyLeuLeuPro 283
6899 GCGTACATTAACACAGCCGCTAGCAAGAGTACCCCTGATGATTGCTTA 6948
283 hAsnThrProArgAsnSerProGlyValProValLysThrGlyGlnArg 299
6949 CCAACAGCCTTAGAAATCCCGAGGTCTTCTGTTAAAGACAGACAGAGA 6998
300 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAsp 316
6999 CTTTCATCTCTATCCAGGAGCTTCCAAAGCCATCAACTCCAGCAGACC 7048
316 cAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTrp 333
7049 TGATCCCACTTCTCTGTTGGCTTGTCTATCCCTCAGGGCCCTCTATT 7098
333 rYrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArg 349
7099 ATGAGGGATGCTTAAGAGAAATTCATGTGACCAAGAGACATAGA 7148
350 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSe 366
7149 AATCATGTACATGGGGTCCCGAAATAGCTTACCTCATCTGAAGTTTC 7198
366 rglLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeu 383
7199 CCGAAGAGGAGACATGATAGAAAGGTCGCCCATCCACCAACACCTTT 7248
383 ystYrSerThrValValTrpGluGlnAlaSerGluAsnGlnTrpLeuVal 399
7249 GCTATAGTACTGTGTTATGAGCAGGCCCTAGAAATACAGTATTACTA 7298
400 ProGlyTrpAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysVal 416
7299 CCTGTATTACAGGTGCTGGGATCATATCTGGGTTTACCCCTGTGT 7348
416 lSerThrSerValPheAsnGlnSerLysAspPheCysValMetValGln 433
7349 TTCTCTCTAGTCTTCAACCAATCCAAAGATTTGTTGTCATGGTCCAA 7398
433 lValProArgValTrpTrpHisProGluGluValValLeuAspGluTrp 449
7399 TCGTCCCGGAGTGTACTTACCATCCTGAGGAAGTGTCTTGATGAAATAT 7448
450 AspTrpArgTrpAsnArgProLysArgGluProValSerLeuThrLeuAl 466
7449 GACTATCGTATPAACCGACCAAAAGAGAACCCGATCTTACCTTAGC 7498
466 aValMetLeuGlnLysGlyThrAlaValGlyValGlyThrGlyThrAla 483
7499 TGTATGCTCGGATTAGGAGCGCGCTTGCGGTAGGAACAGGACAGCTGT 7548
483 lAlaLeuIleThrGlyProGlnGlnLysGlyLysGlyLeuHis 499
7549 CCGTATACAGAGACACAGACAGCTAGGAAGAGCTGGGACTTCAT 7598
500 AlaAlaMetThrGluAspLeuArgAlaLeuLysGlnSerValSerAsnLe 516
7599 GCGGCATGACAGAAAGATCTCCGAGCCTTAGAGAGTCTGTACCAACCT 7648
516 uGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 533
7649 AGAAGAGTCCCTGACTCTTGTGTGAAGTGTCTTCAACAACCGGAGGG 7698
533 lLeuAspLeuLeuPheLeuArgGluGlyLysCysAlaAlaLeuLys 549
7699 GATTAGATCTGCTCTTCTTAAGAGAGTGGTTATGTCCAGCTTAAAA 7748
550 GluGluCysCysPheTrpValAspHisSerGlyAlaIleArgAspSerMe 566
7749 GAAGAATGTGCTCTTATGATCACTCAGAGACCATCAGAGACTCCAT 7798
566 tAsnLysLeuAlaArgLysLysLeuArgArgArgGluArgGluAla 583
7799 GAGCAAGCTTAGAAGAAAGTTAGAGAGCGCTCAAGAGGAAAGAGGCTG 7848
583 spGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThrThr 599
7849 ACCAGGGGTGGTTAAAGATGTTCAACAGGCTCTCTTGGATGACACACC 7898
600 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuThr 616
7899 CTGCTTCTCTGCTGAGCGGAGCCCTAGTAGTCTGCTCTGTTACTTAC 7948
616 rValGlyProCysLeuLeuAsnArgPheValAlaPheValArgGluArg 633
7949 AGTTGGGCTCTCTTAATTATAGTTTGTGCTTGTGTAGAGAAAGAG 7998
633 alSerAlaValGlnIleMetValLeuArgGlnGlnTrpGlnGlyLeuLeu 649
7999 TGAATGAGTCCAGATCATGTTACTTAGGCAACAGTACCAAGGCTTCTG 8048
650 SerGlnGlyLysThrAspLeu 656
8049 AGCCAAGAGAAACTGACCTC 8069

```

```

LMALVTLALAEKSS.INITYDTSRYAFATAHVGAIVYKORGLDITSGAEIKKKEEITSL
LEALHLPKRLAIIPHCPOHOKALIRSGNOMADRYAKOAOAVNLLPIETPKAPDEP
KOYLTPEMOEIKKIDPOESPETGCTGSKETILPHKGLRYEOVIRHRLTGLKHLQ
OLVATSTCYHYLRILPGYADSVYKHCYPCOLYNNPNRPISPKRLGSHGTHGAEWDFTE
VKRPMTAKTYLLVDFVDFSGWBAIFPKKEISTVYAKALLEIIPRREGIPKVISDNG
PAFAVYASOGILAKITLIGIDMKLHCAYRPOSSGOVERNMTIKETLTKLTTEGIDMWIA
LPLPFLAVRNPTEGQGLIPYELLYAGPPLEIYASVADVLLSOPLESLKALEWV
ROARMKDLFEAYSGEDLQVHFHFQGVDSVYRRRAAGNELFTMKGPYLLVLTPTPAV
KVEEISWITHTSHVKLRPLPDSGCMRAEKEENPKLRLHRLVYPSNNNSPCOIHPTLSM
RHLPTRGCEPKRLRILPLSFASITAMFLTLITPQASSKRLIDSSNHPRLSLTWLIDP
DHGTVAVSTGKVPAPRGTMWPELHFLCRLINPARKSTPPNLYRSVGFYOCPEKEKYC
GDSESESTRMSCVTSDNDGDKMPISLDYRKFSTVNSGPKYKMYLKHDKSCSPSD
LDYKIAETFKRKQENIOKWINSMWGITVYRFGGAGSTGLIRLLETGETEPYAVTG
PDKYLAIOGPPALPEPHNLPLVQMLTSLRDIYQSSNGSTGLIPNPNRNSPGVAVT
GORLESLIOGAFOALNSTPDATESSCMLDLSGSPYEGAMKEGFVNTKSHRNOCTM
GSRMKLITLFEVSGKQICIGAPSPHOHLCTSYVVEYEOASENOYLVPGYNPMMACTGL
TPCVSTSVFNQSDPCVMNOIVPRVYHVEEVLEDEYQRYNRPKREVSLLTAVMWL
LGRVAVGTETALITTPGPOLEKGLGEHLEAMTEPDLRLFEVSMLPESTLSLVEVL
ONRGLDITLPLREGGICALKEECCTYVHSCAIDRMSKLRERLERRRRREADQCM
PEGFENFNSPMWMTLLSLTGLTVLVLILLTVGPLINRPVAFRRVSAVOIMVLQOO
YQGLSGOGETDL"
mat_peptide
/gene="pol"
2729..6313
/function="retroviral protease and polymerase"
/product="pol protein"
/gene="pol"
2729..6313
/gene="pol"
6189..8162
/gene="env"
/function="retroviral envelope"
/product="Env protein"
6189..8162
/gene="env"
6812..8918
8811..8816
polysignal
BASE COUNT 2327 a 2245 c 2334 g 2012 t
ORIGIN
alignment_scores:
Quality: 3414.00 Length: 658
Ratio: 5.244 Gaps: 2
Percent Similarity: 98.936 Percent Identity: 97.720
alignment_block:
US-09-171-553B-6 x PEN133816
Align seg 1/1 to: PEN133816 from: 1 to: 8918
1 MethISProThrLeuSerArgAlaGlnHisLeuProThrArgGlyGlyLupr 17
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6189 ATGATGTCGCCAGGTTAAAGCTGGCCGCCACCTCCGAGCTCGGGGAGGAGACC 6238
17 olysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuT 34
6239 GAAAAAGCTGAGATCCCTTAAGCTCCCTCCATCGGCTGGTGTCTTAA 6288
34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50
6289 CTGTAAACAATTAAGTCCCGACGAGCCAGTAGTAAGACGCTTATAGACAGCTG 6338
51 AsnProHisArgProLeuSerLeuThrTrpPheIleIleAspProAspTh 67
6339 AACCOCCTAGACGCTTATCCCTTACCCTGGCTGATTAATGACCCGATAC 6388
67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrpT 84
6389 GGGGTGCTACTGATAATAGCACTCGAGGTGTGTCTCTTAAGGCACACTGGT 6438
84 rProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
6439 GGCCGTGAAGTGCATTTCTGCTCCGATGTGATTAACCCCGCTGTAAAGC 6488

```

```

101 ThrProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117
|||||
6489 ACACCTCCCAACCTAGTCGATGATGGGTCTATGCTGCTGCCAGGAC 6538
117 rGluGlySerGlyTyrGlySerGlyGlySerPheCysArgArg 134
|||||
6539 AGAGAAAGAGAAATCTGTGGGATCTGGGAAATCTCTGTAGAGAT 6588
134 rPseCysValThrSerAsnAspGlyAspTyrPlyThrProIleSerLeu 150
|||||
6589 GGAAGCTGGCTCACCTCCACACATGAGACTGGAATGGCCCATCTCTC 6638
151 GluAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyr 167
|||||
6639 CAGGACCGGGTAAATTCCTCTTGTCAATTCGGCCGGGCAAGTAA 6688
167 sMetLeuLysLeuTyrLysAspLysSerCysSerProSerAspLeuApt 184
|:::|
6689 AGTGATGAATACTATTAAGATTAAGAGCTGCTCCCATCAGACTAGATT 6738
184 rLeuLysIleSerPheThrGlu..ArgLysThrGlyLysTyrSerLysV 200
|||||
6739 ATCTAAGATTAAGTTTCACTGAAAAAGAAAAAGAAATATTTCAAAAG 6788
200 aLAspLysTyrPyrGluLeuGlyAsnSerPheLeu..LeuTyrGlyGly 216
|||||
6789 TGGATTAATGGTATGAGCTGGGGAATAGTTTATTAATATGGCGGG 6838
216 yAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGlu 233
|||||
6839 ACCAGGGTCCACTTACCATTCCGCTTAGAGTACAGACGGGAGACAGAC 6888
233 rOProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProPro 249
|||||
6889 CCCCTGGGCACTGGGAGCCGATTAAGTACTGCTGAACAGGGGCCGCC 6938
250 aAlaLeuGluProProHisAsnLeuProValProGluLeuThrSerLeuAr 266
|||||
6939 GCCCTGAGACCCACGCATTAACCTTGGCGGGGCCCAATTACCTGCTCG 6988
266 gProkAspIleThrGlnProProSerAsnSerThrThrGlyLeuIleProT 283
|||||
6989 GCCTACATAACACACAGTCCCTAGCACAGGTAACCACTGATTGATCTTA 7038
283 hAsnThrProArgAsnSerProGlyValProValLysThrGlyGlnArg 299
|||||
7039 CCAACAGCCCTAGAAACTCCAGAGTGTCTCTTAAAGACAGGACAGAGA 7088
300 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspPr 316
|||||
7089 CTCTTCAGTCTCATCCAGGAGCTTCCAGGCATCAACTCCAGCCGACCC 7138
316 oAspAlaThrSerSerCysTyrPleuCysLeuSerSerGlyProProTyrT 333
|||||
7139 TGATGCCACTCTCTCTGTGGCTTTGTCTATCCCAAGGGCTCTCTTATT 7188
333 yrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArg 349
|||||
7189 ATGAGGGATGGCTAAAGAGAGAAATTCATGTGACCAAGAGCATGAGA 7238
350 AsnGlnCysThrTyrGlySerArgAsnLysLeuThrLeuThrGluValSe 366
|||||
7239 AATCAATGTACATGGGGGTCCGAAATTAAGCTTACCCTCAGCAAGTTTC 7288
366 rGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeuc 383
|||||
7289 CGGGAAGGGGACATGCATAGAAAGCTCCCCATCCACACACACCTTT 7338
383 yTyrSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeuVal 399
|||||
7339 CCTATAGTACTGTGTATATGAGCGGCTCGAATAATCATATTATTA 7388
400 ProGlyTyrAsnArgTyrTyrPalaCysAsnThrGlyLeuThrProCysVa 416

```

```

|||||
7389 CTTGTTATTAACAGGTGGTGGCATGCATATAGGGGTAAACCCCTGTGT 7438
416 lSerThrSerValPheAsnGlnSerLysAspPheCysValMetValGln 433
|||||
7439 TTCACCTCAGTCTTCAACCAATCCAAAGATTTCTGTGTATGGTCCAAA 7488
433 lValProArgValTyrTyrHisProGluGluValValLeuAspGlyTyr 449
|||||
7489 TCGTCCCCCAGTGTACTTACCATCTGAGAAAGTGGCTTGTATGAATAT 7538
450 AspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAl 466
|||||
7539 GACTATGGTATAACGACCAAAAGAGAACCCGTATCCCTACCCAGAC 7588
466 aValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThAla 483
|||||
7589 TGATATGCTCGATTAGGAGCGCGCTTGGCTAGGAACAGGAGACAGCTG 7638
483 lAlaLeuThrGlyProGlnGlnLeuGluLysGlyLeuGluHis 499
|||||
7639 CCTGATCAGACAGACACACAGAGCTAGAGAAAGACTTGGTGAAGTAT 7688
500 AlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnLe 516
|||||
7689 GCGGCATGACAGAGAGATCTCCGAGCTTAGAGAGAGTCTGTAGCAACT 7738
516 uGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 533
|||||
7739 AGAAGAGTCCCTGACTCTTGTCTTGAAGTGTCTTACACACCGGAGGG 7788
533 lLysAspLeuLeuPheLeuArgGluGlyLeuCysAlaAlaLeuLys 549
|||||
7789 GATTAGATCGTGTCTTCTAAGAGAGGTGGTATATGACGCTTAAAA 7838
550 GluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMe 566
|||||
7839 GAAGAAATGTGCTCTTATGATGACTCAGAGGCCATCAGAGACTCCAT 7888
566 tAsnIleLeuLeuArgLysLysLeuGluArgArgArgGluArgGluAla 583
|:::|
7889 GAGCAACCTTAGAGAAAGTTAGAGAGCGCTCGAAGGAAAGAGAGCTG 7938
583 spGlnGlyTyrPheGlnGlyTyrPheAsnArgSerProTyrPheThrThr 599
|||||
7939 ACCAGGGGTGGTTGAAGGATGTTCAACAGGTCTCTTGGATGACACAC 7988
600 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuThr 616
|||||
7989 CTGCTTCTGCTGACGAGACCCCTAGTACTCTGCTCTGTTACTTAC 8038
616 rValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgV 633
8039 AGTTGGGCTCTTCTTAATTAATAGTTTGTGCTTGTAGAGAACAGAG 8088
633 aLSerAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeuLeu 649
|||||
8089 TGAATGACAGTCCAGATCATGTGTAAGTATGAGCAACAGTACCAAGGCTTCTG 8138
650 SerGlnGlyGluThrAspLeu 656
|||||
8139 AGCCAGAGAGAAACTGACCTC 8159

```

seq_name: gb_v1:PEMZ88588

seq_documentation_block: 2000 bp RNA VRL 29-AUG-2000

LOCUS PEMZ88588

DEFINITION Porcine endogenous retrovirus PERV-AB7 env gene for envelope protein, genomic RNA.

ACCESSION AJ288588

VERSION AJ288588.1 GI:7414445

KEYWORDS env gene; envelope protein; porcine endogenous retrovirus.

ORGANISM porcine endogenous retrovirus
 Viruses; Retroviridae; Mammalian type C
 retroviruses; 1-Mammalian type C virus group.

REFERENCE
 1 (bases 1 to 2000)
 Bosch, S., Arnould, C. and Jestin, A.
 Study of full-length porcine endogenous retrovirus genomes with
 envelope gene polymorphism in a specific-pathogen-free large white
 swine herd

JOURNAL
 J. Virol. 74 (18), 8575-8581 (2000)

MEDLINE
 20411432

REFERENCE
 2 (bases 1 to 2000)
 Bosch, S.
 Direct Submission
 Submitted (28-MAR-2000) Bosch S., Molecular Biology Department,
 AFSSA-Zoopole, Zoopole Les Croix B.P.53, Ploufragan 22440, FRANCE

TITLE
 JOURNAL

FEATURES
 source
 1..2000
 /organism="porcine endogenous retrovirus"
 /specific_host="Sus scrofa"
 /db_xref="taxon:61673"
 /clone="PERV-AB7"
 /note="French Large White SPF-pig"
 1..606
 /gene="env"
 1..606
 /gene="env"
 /note="truncated, not functional"
 /codon_start=1
 /product="envelope protein"
 /protein_id="CA86207.1"
 /db_xref="GI:741446"
 /translation="MHPTLSMRHLPRGGEPRRLRILSPASIAWFLTLITPOASSK
 RLIDSSNHRPLSLWLIIDPTGVNSTRGVAPEGTWMPLELHPCRLINPAVKSTP
 PNIVRSYGFCCPGTEKEKYGSGSEFCVRSNNGDMKWPISLDQVKFSPVN
 SGFKYKVMKLYDKSGSPSDLDYLIKISFTEKGNRIIFNSG"

BASE COUNT
 532 a 496 c 499 g 473 t

ORIGIN

alignment_scores:
 Quality: 3317.00 Length: 647
 Ratio: 5.199 Gaps: 3
 Percent Similarity: 98.609 Percent Identity: 96.291

alignment_block:
 us-09-171-553b-6 x PEN288588 ..

Align seg 1/1 to: PEN288588 from: 1 to: 2000

```

1 MethisProthrLeuSerArgArgHisLeuProthrArgGlyGlyLupr 17
|||||
1 ATGCATCCACGTTAAAGTCGGCCACCTCCGAGCTCGGGTGAGAGCC 50
17 olySarGleuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeu 34
|||||
51 GAAAGGCGTGAAGATCCCTTAAAGCTTCGCCATCGCTGGTCCCTTA 100
34 hrLeuThrIleThrProGlnAlaSerSerIlySarGleuIleAspSer 50
|||||
101 CTCTAACAATAACTCCCGAGCGCACTAGTAAAGCCTTATAGACAGCTCG 150
51 AsnProHisArgProLeuSerIleuThrTrpLeuIleIleAspProAsp 67
|||||
151 AACCCCATAGACCTTATCCCTTACCTGCGCTGATTTATGACCTGATAC 200
67 rGlyAlaThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84
|||||
201 GGGTGCTCCTGTAATAGCACTCGAGGTGCTCTTAAGAGCGACCTGGT 250
84 rpProGlnLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
|||||
251 GGCCGGAAGTGCATTTCTGCTCCGATGATTAACCCCGCTGTTAAAGC 300

```

```

101 ThrProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117
|||||
301 ACACCTCCCAACCTAGTCCGAGTATGAGGTCTATGCTGCTGCCAGCAC 350
117 rGlyGlnLeuLysTyrCysGlyGlySerGlyGluSerPheCysArgTr 134
|||||
351 AGAAGAAAGAGAAATAGTGGGGGTCTGAGAAATCCTCTGTAGAGAT 400
134 rpsrCysValThrSerAsnAspGlyAspTrpIlystrProIleSerLeu 150
|||||
401 GGAGCTGCTCCTCCACAGATGAGAGTGAATGCAATGCTCTCTC 450
151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyTyrLy 167
|||||
451 CAGAGCCGGGTAAATATCTCTTGTCAATTCGCGCCGGGCAAGTACA 500
167 SmetLeuLysLeuTyrLysAspLysSerCysSerProSerAspLeuAs 184
|||||
501 AGTATGAACCTATATTAAGATAAGAGCTGCTCCCATCAGACTTAGAT 550
184 yrlLeuLysIleSerPheThrGluArgLysThrGlyLysTyrSerLys 200
|||||
551 ATCTAAAGATAAGTTTCACGTAAGAAAGCAAGAAATATTCAAAGT 600
200 LaspLysTrpTyrGlnLeuGlnLysSerPheLeuLeuTyrGlyGly 216
|||||
601 GGATTAATGGATGAGCTGGGGAATAGTTTATTAATATGCGGGGGA 650
217 AlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGlu 233
|||||
651 ACAGGGTCCACTTAAACATTCGCTTACGATGAGACAGCGGAGAGACC 700
233 opProValAlaMetGlyProAspLysValLeuAlaGlnGlnGlyProPro 250
|||||
701 CCTGTGCGAGTGGAGCCCAATAAGTACTGCTGAACAGGGGCCCCCG 750
250 lalLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeu 266
|||||
751 CCTGGAACCAACCGCATACCTGCGGGTGCCCAATTAACCTGCTGGG 800
267 ProAspIleThrGlnProProSerAsnSerThrThrGlyLeuLeuPro 283
|||||
801 CCTACATACACACGCGCTAGCAAGGATGACGATGATGATTCCTAC 850
283 rAsnThrProArgAsnSerProGlyValProValLysThrGlyGlnArg 300
|||||
851 CAACAGCCCTAGAAACTCCCGAGTGTCTCTGTTAAACAGACAGACAGA 900
300 eupHiserLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAsp 316
|||||
901 TCTTCACTCATTCACAGGAGGCTTCCAGACCAACATCCACGAGACC 950
316 oAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTyr 333
|||||
951 TGATGCCACTTCCTTGTGGCTTGTCTATCTCCTCAGGGCCCTTATT 1000
333 yrlGlnGlyMetAlaLysGlnArgLysPheAsnValThrLysGlnHis 349
|||||
1001 ATGAGGGGATGCTAAAGAAAGAAATTCAAATGACCAAGACAGCATAGA 1050
1001 ATGCAATAGTATGAGGGGCTCCGAAATAATACCTTACCTGAAATTT 1100
350 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGlnVal 366
|||||
1101 CGGGAAGGAGCATGCAATAGAAAGCACTCCCATCCACCAACACTTT 1150
366 rGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeu 383
|||||
1151 GCTATATCTGCTGCTTATGACAGAGCCGCAAGAAATCAGTATTTAGTA 1200
383 ystTyrSerThrValIleValTyrGlnAlaSerGlnLysGlnTyrLeuVal 399
|||||
1200 ProGlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysVal 416

```



```

1201 CCTGGTTGAACAGGTGGGCGATGCAATACGTGCTTACCCCTGTGT 1250
416 IserThrSerValPheAsnGlnSerLysAspPheCysValMetValGlnI 433
1251 PTCACCTCAGCTTCACCAATCCAAAGATTTCGTGTCAATGAGTCCAAA 1300
433 IValProArgValTyrThrHisProGluGluValValLeuAspGluTyr 449
1301 TCGTCCCGGAGTGTACTACATCCGAGAGAGTGTCTCTTCAATAT 1350
450 AspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAl 466
1351 GACTATCGGTATTAACGACCAAAAGAGACCCGTATCCCTTACCTAGC 1400
466 aValMetLeuGlyLeuGlyThrAlaValAlaGlyAlaGlyThrGlyPheAla 483
1401 TGTAAATGCTCGATTAGGAGCGCGCTGTGCTAGAACAGGACAGCTG 1450
483 IAlaLeuIleThrGlyProGlnGlnLeuGlyGlyLeuGlyGluLeuHis 499
1451 CCGTATATCAGAGACACAGCAGCTAGAGAAAGACTGTGTAGCTACAT 1500
500 AlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnLe 516
1501 CGGTCATGACAGAAAGATCTCCGACCTTAGAGAGCTGTGTAGCAACT 1550
516 uGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 533
1551 AAGAGAGTCCCTGACTCTTGTCTGAAGTGGTCTTACAGAACCGGAGGG 1600
533 ILeuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaAlaLeuLys 549
1601 GATTAGATCTGCTGTTTCTAAGAGAGTGGTGTATGTCCACCTTAAAA 1650
550 GluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerLe 566
1651 GAAGATGTTGCTTATATGATCATCATCAGAGACCATCAGACTCAT 1700
566 tAsnLysLeuArgLysLysLeuGluArgArgArgGluArgGluAla 583
1701 GAGCAAGCTTAAAGAAAGTTAGAGAGCGCTGAAAGGAAAGAGGCTG 1750
583 spGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThrThr 599
1751 ACCAGGGGTGTTTGAAGATGTTTCAACAGCTTCCTTGGATGACCAAC 1800
600 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeu 616
1801 CTGCTTTCGTCTGACGGGACCCCTAGTAGTCTGCTCTCTTACTTAC 1850
616 tValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArg 633
1851 AGTGGGCTTGTCTTAATTAATGAGTTGTGCTTGTGTAGAACGAG 1900
633 aIserAlaValGlnIleMetValLeuArgGlnIle 644
1901 TAGGTGACATCCAGATCATGTACTTAGCAACGC 1935
seq_name: gb_v1: PEN288589
seq_documentation_block:
LOCUS PEN288589 1968 bp RNA VRL 29-AUG-2000
DEFINITION Porcine endogenous retrovirus PERV-B1 env gene for envelope
protein, genomic RNA.
ACCESSION AJ288589
VERSION AJ288589.1 GI:7414447
KEYWORDS env gene; envelope protein.
SOURCE porcine endogenous retrovirus.
ORGANISM Porcine endogenous retrovirus.
Virus; Retroviral; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
REFERENCE 1 (bases 1 to 1968)

```

```

AUTHORS Bosch, S., Arnould, C. and Jestin, A.
TITLE Study of full-length porcine endogenous retrovirus genomes with
envelope gene polymorphism in a specific-pathogen-free large white
swine herd
JOURNAL J. Virol. 74 (18), 8575-8581 (2000)
MEDLINE 20411432
REFERENCE 2 (bases 1 to 1968)
AUTHORS Bosch, S.
TITLE Direct Submission
JOURNAL Submitted (28-MAR-2000) Bosch S., Molecular Biology Department,
AFSSA-Zoopole, Zoopole les Croix B.P.53, Ploufragan 22440, FRANCE
FEATURES
source
1..1968
/organism="porcine endogenous retrovirus"
/specific_host="Sus scrofa"
/db_xref="taxon:61673"
/clone="PERV-B1"
/note="French Large White SPF-pig"
1..1968
/gene="env"
1..1968
/gene="env"
CDS
1..1968
/codon_start=1
/product="envelope protein"
/protein_id="CAB86208.1"
/db_xref="GI:7414448"
/translation="MHPTLSRHPPTRGGEPRKLGILPFSGIAMFLITLITTPQASAK
RLIDSSNPRLSLMLIIDPTGVTVNSRVARGRMTMMPLHLCLRLINPAVSTP
PMLVRSYGFCCGPGMEKEKESGSGESGRMSCTVSDGDKKMFISLDIPKFSFN
SGPKKRWKRLKDKSCSPSDIYKISTEYGRKBNOKMNGMSKIVKYKCGA
GSTLTIIENGTEPVAAGVDKVLADGPPALBPPLNLPOLTSLRPTDPSNGTT
GLIPNTPHSHSGVPVKTGQRLFLSLQGFQAINSTDPAISSCLSSGPPYEGM
AKEGKFNVIKERNOCITWSRNKLTLEVSGKTCIGRAPSHQHCSTVAYEBOASE
NOYLVDGYRRWMAWNGTLPCVSTSVFNOSKDFCAGAPVPRVYVHPEMVDYEYR
VNRPKREPSTLIAYMLGIGTAVGVGTGTAAXIITGPQOLEKGLPHHAMEDLRALE
EVSNSLESITLSLEVYLDNRKGDILPLIREGLDALKEBCECFVDSGAIROMSK
LRLRLERRREREADQGWFEWPNRSPMWTLLSLALTEPLVLLVLLITVGPOLNRFV
AEVRERVSAAVOIMVLRORLMKGEFQHTGGRY"
BASE COUNT 522 a 483 c 492 g 468 t 3 others
ORIGIN
alignment_scores:
Quality: 3288.50 Length: 646
Ratio: 5.187 Gaps: 3
Percent Similarity: 98.142 Percent Identity: 95.820
alignment_block:
US-09-171-553B-6 x PEN288589 ..
Align seg 1/1 to: PEN288589 from: 1 to: 1968
1 MethisProThrLeuSerArgArgHisLeuProThrArgGlyGluIupr 17
|||||
1 ATGATCCACAGCTTAAGCTGGGCGCAACCCCGACTGGGGGGGAGACC 50
17 OlyArgLeuArgGlyLeuProLeuSerPheAlaSerIleAlaIleTrpPheLeu 34
|||||
51 CAAGAAGACTGGGCAATCCCTTAACCTCGGCTCATCGCTGCTCTCTTA 100
34 hLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50
|||||
101 CTCTTACATTAACCTCCAGCGCAGTAGTAACCCCTTATAGACAGCTCG 150
51 AsnProHisArgProLeuSerLeuThrTrpLeuIleIleAspProAspTh 67
|||||
151 AACCCCATAGACCTTTATCCCTTAACCTGGCTGATATGACCTGATAC 200
67 rGlyValThrValAsnSerThrArgGlyValAlaIleProArgGlyThrTrp 84
|||||
201 GGGGTGTACTGTAAATAGACACTCGGGGTGTGCTCTAGAGGACACCTGGT 250
84 rProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100

```

```

|||||
251 GGCCTAAGCATTTCTGCTCCGATTAATTAACCCGCTGTAAAC 300
101 TheProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117
|||||
301 AACCTCCCACTAGTCCGAGTATGGGTCTATGCTGCCAGGCAC 350
117 rGlyuSgLuYsTyrCysGlySerGlyuSerPheCysArgArgT 134
|||||
351 AGGAGAAAGAAATACTGTGGGGGTTCTGGGAATCTCTAGAGAT 400
134 rSerCysValThrSerAsnAspGlyAspTyrTrpProIleSerLeu 150
|||||
401 GGAGCTGCGCTCACTCAAGAGAGAGAGTGAATGGCCGATCTCTGC 450
151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyr 167
|||||
451 CAGGACCGGGTAAATTTCTCTTGTCAATTCGGGCGCGGCAAGTACA 500
167 SMetMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAspT 184
|||||
501 AGTGATGAACATAATATAAGATAGAGCTGCCATCAGACTTAGATT 550
184 YrLeuYsLleSerPheThrGlu..ArgLysThrGlyLysTyrSerLysV 200
|||||
551 ATCTAAGATTAAGTTTCACTGAAGAAAGAAAGAAATATGCATAAG 600
200 aAspLysTrpTyrGluLeuGlyAsnSerPheLeu..LeuTyrGlyGly 216
|||||
601 TGGATGAATGATATGAGCTGGGGAATAGTTTATTAATATAGGGGGG 650
216 yAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGluP 233
|||||
651 ACCAGGGTCCACTTAACCATTT.....ATAGAGAGCGGAGACAGAC 691
233 rOProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProPro 249
|||||
692 CCCCTGTGGCAGTGGAGCCCGATTAAGTACTGGCTGAACAGGGGCCCCG 741
250 AAlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuAr 266
|||||
742 GCCCTGGAGCCACCGCATTAACCTGGCGGTCGCCCAATTAACTCCCTGG 791
266 gProAspLleThrGlnProProSerAsnSerThrThrGlyLeuIleProT 283
|||||
792 GCTTCACATTAACACAGCGCTTGAACAGCGTCACTGATTTGATTCCTA 841
283 hAsnThrProArgAsnSerProGlyValProValLysThrGlyGlnArg 299
|||||
842 CCAACACGCGCTAGACACTCCCAAGGTGTTCTGTTAAGCAGGACAGAGA 891
300 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspPr 316
|||||
892 CTCCTCAGTCTCATCCAGGAGCTTCCCAAGCATCACTCCACCGACCC 941
316 oAspAlaThrSerSerCysTyrPheCysLeuSerSerGlyProProTyrT 333
|||||
942 TGAATCCCACTCTCTGTTGCTTGTCTATCCCTCAGGGGCTCTCTATT 991
333 yfGluGlyMetAlaLysGluArgLysPheAsnValThrLysGlnHisArg 349
|||||
992 ATGAGGGGATGGCTAAAGAGAAATCAATGATGACCAAGAGAGATAGA 1041
350 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSe 366
|||||
1042 AATCATATGAGTGGGGGTCGGAATATAGCTTACCTCACTGAAGTTTC 1091
366 rGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeuc 383
|||||
1092 CGGGAAGGAGCATGATAGGAAAGCTCCCATCCACCAACACCTTT 1141
383 yTyrSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeuVal 399
|||||

```

```

1142 GCTATAGTACTGTANTTTATGAGCAGGCTTCAGAAATACATATTAGTA 1191
400 ProGlyTyrAsnArgTyrTrpAlaCysAsnThrGlyLeuThrProCysVa 416
|||||
1192 CTTGTTTATACAGTGGTGGGATGCATATCTGGGTTAACCCCTGTCT 1241
416 lSerThrSerValPheAsnGlnSerLysAspPheCysValMetValGlnI 433
|||||
1242 TTCCACCTCACTTCAACCAATCCAAAGATTTCTGTGTCATGCTCCAA 1291
433 lValProArgValTyrTyrHisProGluGluValIleLeuAspGluTyr 449
|||||
1292 TCGTCCCGCGAGTACTACCATCTTAAGAAATGTCCTTGATGAATAT 1341
450 AspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAl 466
|||||
1342 GAATATCGGTATACCGACCAAAAGAAAGAACCCGTATCCCTAACCATAGC 1391
466 aValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAla 483
|||||
1392 TGTATGCTCGATTAGGAGCGCGCTTGGCGTAGAAACAGGACAGCTG 1441
483 lAlaLeuIleThrGlyProGlnGluLeuGlyLysGlyLeuGluLeuHis 499
|||||
1442 CCMTGATCACAGGACACAGCAGTTAGAGAAAGGACTTGTGAGCTTCAT 1491
500 AAlaIleThrGluAspLeuArgAlaLeuLysGlnSerValSerAsnLe 516
|||||
1492 GCGGCATGACAGAAAGATCTCCGAGCTTAGAGAGAGCTGTAGCAACCT 1541
516 uGluGlnSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgG 533
|||||
1542 AGAAGAGTCCCTGACTCTTGTCTGAAGTGCTTCAACAACCGGAGG 1591
533 lYleuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaIleLeuLys 549
|||||
1592 GATTAGATCTGCTTTCTTAAGAGAGAGTGGGTGTGTGACCTTAA 1641
550 GluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMe 566
|||||
1642 GAAGAATGTTGCTCTTAATGATCACTCAAGAGCCATCAAGAGCTCAT 1691
566 tAsnLysLeuArgLysLysLeuGluArgArgArgGluArgGluAla 583
|||||
1692 GAGCAACCTTAGAGAAAGGTTAGAGAGCGCTCGAAGGAAAGAGAGCTG 1741
583 sPGLnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThrThr 599
|||||
1742 ACCAGGGGTGTTGAAGAGATGTTCAACAGGTCTCTTGATGATCACCC 1791
600 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuThr 616
|||||
1792 CTGCTTCTGCTGAGAGGAGCCCTAGTAGTCTGCTGCTGTTACTTACT 1841
616 rValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgV 633
|||||
1842 AGTTGGGCTTGTCTTAATTAATAGGTTTGTGCTTGTAGAGAACGAG 1891
633 aLserAlaValGlnIleMetValLeuArgGlnGln 644
|||||
1892 TGAATGACATCAGATCATGTTACTTAGACACAGC 1926
seq_name: gb_v1.PEN288592
seq_documentation_block: 2000 bp RNA VRL 29-AUG-2000
LOCUS PEN288592
DEFINITION Porcine endogenous retrovirus PERV-B7 env gene for envelope
protein, genomic RNA.
ACCESSION AJ288592
VERSION AJ288592.1 GI:7414453
KEYWORDS env gene; envelope protein.
SOURCE porcine endogenous retrovirus.
ORGANISM porcine endogenous retrovirus

```

Viruses: Retrovirdae: Retroviridae: Mammalian type C
 retroviruses: 1-Mammalian type C virus group.
 REFERENCE
 1 (bases 1 to 2000)
 Bosch, S., Arnould, C. and Jestin, A.
 Study of full-length porcine endogenous retrovirus genomes with
 envelope gene polymorphism in a specific-pathogen-free large white
 swine herd
 JOURNAL
 1. Virol. 74 (18), 8575-8581 (2000)
 MEDLINE
 20411432
 2 (bases 1 to 2000)
 Bosch, S., Arnould, C. and Jestin, A.
 Study of full-length porcine endogenous retroviral (PERV) genomes
 with envelope gene polymorphism in a large white
 specific-pathogen-free swine herd
 JOURNAL
 3 (bases 1 to 2000)
 Unpublished
 REFERENCE
 Bosch, S.
 Direct Submission
 Submitted (28-MAR-2000) Bosch S., Molecular Biology Department,
 AFSSA-zoopole, Zoopole Les Croix B.P.53, Ploufragan 22440, FRANCE
 JOURNAL
 FEATURES
 source
 1. .2000
 /organism="porcine endogenous retrovirus"
 /specific_host="Sus scrofa"
 /db_xref="taxon:61673"
 /clone="PERV-B7"
 /note="French Large White SPF-pig"
 1. .1899
 /gene="env"
 1. .1899
 /gene="env"
 /codon_start=1
 /product="envelope protein"
 /protein_id="CAB86211.1"
 /db_xref="GI:741454"
 /translation="MHPILSWRLPPTGGERKRLIPLSFASFAMLTLTITIPQASSK
 RLIDSSNPHRLSLTLLIDPDTGVNNTGVAIPRGTWPELHFLRLINPAVKSP
 PNLVRSYFCPCPEKEKCYGSGSEFCRMSCVTENDGMKPIFLORVKEFVN
 SGPGRYVMKLYKDKSCSPDLVLTISFEKGEKQENIKMNGIYEVKGGGA
 GSTVTRIRITETGEPYAVSPDKVLAEOGPALPEPHNI.PVRLTSLRDIDQPPN
 GTGLIPTNMRNSPVPVKTGKRLFLSLIGAFKALSTIDPEANSSCWLCISGPPY
 EGMKEGFVNTKEHKNCTWGSNNKLTLEVSQKGTICGAPSHOPECTSTVYED
 ASENQYLPVGNWACNTGLPCVSTSVFNQSDPCVMOIVPVYVYHBEVLEVDY
 DYRNRPKEGPVSLTAVMLGLTAVGVTGTAITLITGPOOLEKGLGELHAAATEDR
 ALESVSNLEESLTLSGVVONRGLDLPLRGGICAALEKCCFYVDSGAIKRS
 MSKLRERLERRRRREADQGWFEQEMNRPMTLLSALGPLVVLILLITVGCILIN
 KFAVARE"
 BASE COUNT 532 a 485 c 505 g 477 t 1 others
 ORIGIN
 alignment_scores:
 Quality: 3278.00 Length: 646
 Ratio: 5.179 Gaps: 2
 Percent Similarity: 97.988 Percent Identity: 95.511
 alignment_block:
 US-09-171-553B-6 x PEN288592 ..
 Align seg 1/1 to: PEN288592 from: 1 to: 2000
 1 MethisProthleuSerArgHisleuProThrArgglyIupr 17
 1 ATGCATCCCAAGTTAAGCTGGCGCACCTCCGACCTGGGGTGAGAGCG 50
 17 oLysArgLeuArgIleProleuSerPheAlaSerIleAlaTrpPheLeu 34
 51 GAAAAGACTGAGAAATCCCTTAACCTTCGCTGCTGCTGCTGCTTA 100
 34 hrlaThrIleThrProGlnAlaSerIleArgLeuIleAspSerSer 50
 101 CTCTAACAAATATTCCTCCAGGCCAGTAGTAACGCTTATAGACAGCTCG 150

51 AsnProHisArgProLeuSerIleuThrTrpLeuIleIleAspProAsp 67
 151 AACCCTCCATAGACCTTATCCCTTACCTGAGCTGATTTATGACCTGTATAC 200
 67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84
 201 GGGTGTCTACTGTAAATAGACACTGAGGTGGTGTCTTACAGAGCCCTGGT 250
 84 rPProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
 251 GGCCTGAACCTCATTTCTGCTCCGATTTGATTAAACCCGGCTGTAAGAAGC 300
 101 ThrProAsnLeuValAlaSerTrpGlyPheTrpCysCysProGlyThr 117
 301 AACCTCCCAACCTAGTCCGAGTTATGGCTTATGCTCCAGGACAC 350
 117 rGluGluGluLysTrpCysGlyGlySerGlyGluSerPheCysArgTrp 134
 351 AGAGAAAGAGAAATACCTGGGGGTTCTGGGGGAATCCTGTGAGAGAT 400
 134 rPserCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150
 401 GGAGCTGCTGCTACCTCCACAGCATGAGAGCTGGAATGGCCGATCTTCTC 450
 151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTrp 167
 451 CAGGACCGGTAAATTCCTCTTGTCAATTCGCGCCCGGCAAGTACAG 500
 167 sMetMetLysLeuTrpLysAspLysSerCysSerProSerAspLeu 184
 501 AGTGATGAACCTATATTAAGATAGAGCTGCTCCATCCACACTTATGATT 550
 184 YrLeuLysIleSerPheThrGlu..ArgLysThrGlyLysTrpSerLys 200
 551 ATCTAAGATAGATTTCACCTGAGAAAGAGAAAGAGAAATATTCAGAAAG 600
 200 aLAspLysTrpTrpGlyLeuGlyLysSerPheLeu..LeuTrpGlyGly 216
 601 TGGATTAATGTGATAGCTGGGGAATGTTTTTATAAATTTGGCGGGGG 650
 216 ValGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGlu 233
 651 AGCAGGGTCCACTGATACCATTCGCTTAGATAGACAGGGGACAGAAC 700
 233 rOProValAlaMetGlyProAspLysValLeuAlaGluGluGlyPro 249
 701 CCCCTGTGGCACTGGAGCCCGCATTAAGTACGCTGAACAGGGGCCCG 750
 250 AlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeu 266
 751 GCCCTGGAGCCACCGCATTAATTCGAGTGGCCCAATTAACCTCGCTCG 800
 266 gProAspIleThrGlnProProSerAsnSerThrThrGlyLeuIlePro 283
 801 GCTTCATATACACAGCCCGCTAGCAGCGTACACTGATTGATTCTCTA 850
 283 hTrAsnThrProArgAsnSerProGlyValProValLysThrGlyLys 299
 851 CCAACATGCTAGAAATCCCGAGGTGCTCTGTTAAGACAGAGAAAGAA 900
 300 LeuPheSerLeuIleGluGlyAlaPheGlnAlaIleAsnSerThrAsp 316
 901 GCTTCATGCTCATCCAGGAGCTTTCAGAGCAATCAATTCACCGACCC 950
 316 oAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTrp 333
 951 GAGAGGCAATCTCTTGTGCTGCTTGTCTATCCTCAGGCGCTCTTATT 1000
 333 YrGluGlyMetAlaLysGlyLysPheAsnValThrLysGluLysArg 349
 1001 ATGAGGGATGTGCTAAAGAGAAATTCATGTCACCAAGAGATAGAGA 1050
 350 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrIleThrGluVal 366

```

1051 AATCATATGATGAGGGTCCGAAATTAAGCTTACCTCAGCAAGATTTTC
366 rgllysglythrCysAlleglyLysAlaProProSerHisGlnHisLeuc 383
1101 CGGGAAGGAGCATGATGAGAAAGCTCCCATCCCAACCAACACCTT 1150
383 ystfserthrValValItyrGluGlnAlaSerGluAsnGlnItyrLeuVal 399
1151 GCTATATGACTGTGTTTATGAACAGGCTCAGAAATCATGATTTAGTA 1200
400 ProGlyTyrAsnArGtrPrpAlaCysAsnThrGlyLeuThrProCysVal 416
1201 CCTGGTTATACAGGTGGGCAATGCACTAGGTTAACCCCTGTGT 1250
416 lserThrSerValPheAsnGlnSerLysAspPheCysValMetValGlnI 433
1251 TTCACCTCAGTCTTAAACCAATCCAAAGATTCTGTGTCAGTGCCAAA 1300
433 lvalProArgValItyrTyrHisProGluGlnValValLeuAspGluTyr 449
1301 TCGTCCCGCCGAGTACTACCATCCTGAGGAGTGTCTCTGATGAATAT 1350
450 AspTyrArgTyrAsnArGProLysArgGluProValSerLeuThrLeuAl 466
1351 GACTATCGGTATACCGACCGAAGAGAACCCGTTATCCCTTACCTAGC 1400
466 aValMetLeuGlyLeuGlyThrAlaValAlaGlyValGlyThrAlaAla 483
1401 TGTATATGCTCGGATTAGGAGCGCGCTGGCGTAGAAGAGGACAGCTG 1450
483 lalLeuIleThrGlyProGlnGlnLeuGlyGlyLeuGlyGluLeuHis 499
1451 CCTATATACAGACGACACAGACAGCTGAGAAAGACTGTGACGTACAT 1500
500 AlaAlaMetThrGluAspLeuArGAlaLeuLysGluSerValSerAspLe 516
1501 GCGGCATGACAGAAAGATCTCCGAGCTTACAGAGTGTGTAGCAACT 1550
516 uGluGlnSerLeuThrSerLeuSerGluValValLeuGlnAsnArGArg 533
1551 AGAAGAGTCCCTGACTCTTGTCTGGAGTGTCTACAGAACCGGAGG 1600
533 lylLeuAspLeuLeuPheLeuArgGluGlyLeuCysAlaAlaLeuLys 549
1601 GATTAGATCTCTGTTCTTAAGAGAGGTGGTTATGTGACGCTTAAA 1650
550 GluGluCysCysPheTyrValAspHisSerGlyValAlaIleArgAspSer 566
1651 GAAGATGTCTCTTATGATGATCCTCAGAGCCATCAGAGACTCAT 1700
566 tAsnLysLeuArGlyLysLeuGluArGArgArGluArgGluAla 583
1701 GAGCAAGCTTAGAGAAAGGTTAGAGAGCGCTGAAAGGAAAGAGAGCTG 1750
583 spGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTyrMetThrThr 599
1751 ACCAGGGGTGGTTGAAGGATGTTCAACAGGTCCTTGGATGACACACC 1800
600 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeu 616
1801 CTGCTTCTGCTGACGAGGAGCGCTAGTAGTCTGCTCTGTTACTTAC 1850
616 rValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgy 633
1851 AGTTGGGCTTGTCTTAATTAAGGTTGTTCCTTGTAGAGAAATGAG 1900
633 alSerAlaValGlnIleMetValLeuArgGlnGln 644
1901 TGAGTGCAGTCCAGATCATGTGTTAGACAACGC 1935
seq_name: gb_v1.PEN288590

```

```

seq_documentation_block:
LOCUS      PEN288590      1996 bp      RNA
DEFINITION Porcine endogenous retrovirus PERV-B4 env gene for envelope protein, genomic RNA.
ACCESSION AJ288590.1 GI:7414449
VERSION    AJ288590.1
KEYWORDS   env gene; envelope protein.
SOURCE      porcine endogenous retrovirus.
ORGANISM   porcine endogenous retrovirus.
REFERENCE   1 (bases 1 to 1996)
AUTHORS     Bosch, S., Arnould, C. and Jestin, A.
TITLE       Study of full-length porcine endogenous retrovirus genomes with envelope gene polymorphism in a specific pathogen-free large white swine herd
JOURNAL     J. Virol. 74 (18), 8575-8581 (2000)
MEDLINE     20411432
REFERENCE   2 (bases 1 to 1996)
AUTHORS     Bosch, S.
TITLE       Direct Submission
JOURNAL     Submitted (28-MAR-2000) Bosch S., Molecular Biology Department, AFSSA-Zoopole, Zoopole des Croix B.P.53, Ploufragan 22440, FRANCE
FEATURES
source
1..1996
/organism="porcine endogenous retrovirus"
/specific_host="Sus scrofa"
/db_xref="taxon:61673"
/clone="PERV-B4"
/note="French Large White SPF-pig"
1..1215
/gene="env"
1..1215
/gene="env"
CDS
1..1215
/note="truncated, not functional"
/codon_start=1
/product="envelope protein"
/protein_id="CAB86209.1"
/db_xref="GI:7414450"
/translation="MRRPSSRRHPTKGGEKRLRIPLSFASIAMPLTLVTTPQASKRLISSDPHRLSTLWLIIIDPTVYTNASTGVAPRGTWMELEFLRLINPAKSTPPLNVRSGYCCPTEKEKYCGSGRESFCGMSCVTSNDMDKWPISLDVRKFSFVNSGPEKYKVMELYKDKSCSPDLDYKISFPEKSGKGNLQKWINMSWGIYFKYGGGASTLIRLRIETGHTPVAVGNIVLAEOPPALEPHNIPVPLQTLRDIDPPNSTGILPNTPEKNSGVVVKRGORLFLIGARQALSTPDASSSCWLCSSPPYEGMARREGFNTKEHRQCTMGSNKLTILVBSKGTCTGKAPPSHOLCNSTVIEQALEMNYLPQGIN"
BASE COUNT      521 a      499 c      503 g      472 t      1 others
ORIGIN
alignment_scores:
Quality: 3108.50      Length: 647
Ratio: 5.022      Gaps: 4
Percent Similarity: 95.672      Percent Identity: 91.345
alignment_block:
US-09-171-553b-6 x PEN288590
Align seg 1/1 to: PEN288590 from: 1 to: 1996
1 MetHisProThrLeuSerArGArGHisLeuProThrArGlyGlyLupr 17
||||| 1 ATGCGTCCCAAGCTCAAGCGCGGACCACTCCGAGCTGGGAGAGACC 50
17 olysArgLeuArgGlyLeuProLeuSerPheAlaSerIleAlaTrpPheLeu 34
||||| 51 GAAAGACATGAGAAATCCCTTAAGCTTCCGCTCCATCGCTGTGTTCTTA 100
34 hLeuThrIleThrProGlnAlaSerLysArgLeuIleAspSerSer 50
||||| 101 CTCTAACAGTGTCTCCCAAGGCGCACTAGTAACGCTTATGACAGAGCTCG 150

```

```
51 AsnProHisArgProLeuSerLeuThrTrpLeuIleIleAspProAspThr 67
151 GACCCCTAGACCTTACCTTACGGCGGATTTAGACCCGATAC 200
67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84
201 GGGTGTCACTGTTATAGCACTCGAGGTGCTGCTTAGAGCACCCTGT 250
84 rPProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
251 GGCCTGAACCTGATTTCTCCCTCCATGATTAAACCCGCTGTAAAAAC 300
101 ThrProProAsnLeuValArgSerTrpGlyPheTrpCysCysProGlyThr 117
301 ACACCTCCCAACCTAGTCCGAGTATGGGTTCTATGTCGCCAGGCAC 350
117 rGluLysGluLysTrpCysGlyGlySerGlyLysSerPheCysArgArgT 134
351 AGAGAAAGAGAACTACTGTGGGGGTTCCAGGGAATCTTGTGGGAGAT 400
134 rPserCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150
401 GGAGCTGCGTCACTCCAAAGATGAGACTGGAAATGCCGATCTCTCTC 450
151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTrp 167
451 CMGAGCCGGGTAATAATTCCTCTTGTCAATTCGGCGCCGGGCAAGTACA 500
167 sMetMetLysLeuTrpLysAspLysSerCysSerProSerAspLeuAspT 184
501 AGTATGAGAACTATATAAGATAGAGCTGCTCCCATCATGACTTAGATT 550
184 rLeuLysIleSerPheThrGlu..ArgLysThrGlyLysTrpSerLysV 200
551 ATMTAAAGATAAGTTTCACTGAAAGAAACAGGGAATATATCAAAAG 600
200 alaSplStrTrpTrpGluLeuGlyAsnSerPheLeu..LeuTrpGlyGly 216
601 TGGATTAATGATGATGAGCTGGGGAATGTTTATATAATATGAGAGGG 650
216 yAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGluP 233
651 AGCAGGGTCCACTTAACATATCGCTTAGAGATGAGAGCGGGCAGCGAA 700
233 rProValAlaMetGlyProAspLysValLeuIleGluGlyProPro 249
701 CCCCCTGGCAGTGGGACCCACATAGTACTGGCTGAACAGAGCCCGG 750
250 AlaLeuGluProProHisAsnLeuProValProGluLeuThrSerLeuAr 266
751 GCCCTGGAGCCACACAACTTGGCGGTGCCCACTTAACCTCGCTGGG 800
266 gProAspLysThrGlnProProSerAsnSerThrThrGlyLeuIleProT 283
801 GCCTGACATTAACACAGCGCGCTGACAAACAGTACCTGATGATTCTTA 850
283 hrAsnThrProArgAsnSerProGlyValProValLysThrGlyGlnArg 299
851 CCAACAGCCTAGAACTCCCGGTCTCTCTCTTAAGACAGACAGAGAGA 900
300 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspP 316
901 CTCTTCAGTCTCATCCAGGAGCTTCCAAAGCATCAACTCCACGACCC 950
316 cAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTrpT 333
951 CGATGCCAGTCTTCTGTGGCTTGTCTATCTCAGGGCTCTCTTATT 1000
333 yArgGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArg 349
1001 ATGAGGGAATGCTAGAGAAAGGAATTCATGATGACCAAGAGCATAGA 1050
350 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSe 366

:::
1051 GATCAATGTACATGGGGGCTCCGAAATTAAGCTACCTCACTGAAGTTTC 1100
366 rGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeuc 383
1101 CGGGAAGGGGACATGATAGGAAAGCTCCCATCCACACACACCTTT 1150
383 yTrpSerThrValValTrpGluGlnAlaSerGlnAsnGlnTrpLeuVal 399
1151 GCATAGTACTATGTTTATGAGCAGGCTTAGAAATCAGTATTGTA 1200
400 ProGlyTrpAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysVa 416
1201 CCTGTTATAACTGATGGTGGCATGCAATACGCGTTAATCTCCTGTGT 1250
416 lSerThrSerValPheAsnGlnSerLysAspPheCysValMetValGln 433
1251 TTCCACCAAGCTTTTAACCAATCCAAAGATTCTGTGTCATGTGTCAGC 1300
433 lValProArgValTrpTrpHisProGluGluValValLeuAspGluTrp 449
1301 TCGTCCCTGGGTACACTACCATCTGAGGAAGTACTCTGATGAGTTT 1350
450 AspTrpArgTrpAsnArg..ProLysArgGluProValSerLeuThrLeu 466
1351 GACTATCGCCTACCCTCCAAAGAGACCCCTGACCCCTAACCTAG 1400
466 lValMetLeuGlyLeuGlyThrAlaValAlaGlyValGlyThrGlyAla 482
1401 CCGTTATACGTAGATTAGGATGAGTCT..GGTGGGGAGACAGGACAGCT 1447
483 AlaLeuIleThrGlyProGlnGluLeuGluLysGlyLeuGlyLeuHis 499
1448 GCCCTGATCACAGGACACACAGCAGCTAGAAAGAGCTAGTGGTACACA 1497
499 sAlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnL 516
1498 TGCCGCCATTAACAGAAATCTCCAGCCCTAGAGAAATCTTTAAACC 1547
516 euGluGluSerLeuThrSerLeuSerGluValAlaLeuGlnAsnArgArg 532
1548 TAGAAGAAATCCCTGATTTCTTGTCTGTAAGTGTTTACAACTGGAGG 1597
532 GlyLeuAspLeuLeuPheLeuArgGluGlyLeuCysAlaAlaLeuLys 549
1598 GGATTAGATCTGCTTCTTAAGAAAGAGTGGTATGTGACGCTTAAG 1647
549 sGluGluCysCysPheTrpValAspHisSerGlyAlaIleArgAspSerM 566
1648 ACAGAAATGTGCTTCTATGTGATCACTCAGGAGCCTACAGAGCTCCA 1697
566 eAsnLysLeuArgLysLysLeuGluArgArgArgGluArgGluAla 582
1698 TAGCAGAGCTTAGAAGAAAGTTAGAGGAGCGTCGAAGGAAAGAGAGCT 1747
582 AspGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThrTh 599
1748 GACCAAGGGGTGTTGAAGAGATGTTCAACAGGTCTCTTGGATACCA 1797
599 rLeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuT 616
1798 CCGCTTGTGCTGTGATGGAGCTCTAGTGGTCTCTGCTCTGTTGCTTA 1847
616 hrValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArg 632
1848 CAGTTGGGCTTGCTTAATTATAGTTGTGCTTGTGTAGAGAAGA 1897
632 ValSerAlaValAlaIleMetValLeuArgGlnGln 644
1898 GTGAGTGCAGTCCAGATCATGTGTTACTTACACAA 1933
seq_name: gb_v1:PEN288591
```

LOCUS DEFINITION	seq documentation block:	2002 bp	RNA	VRL	29-AUG-2000
PROTEIN	porcine endogenous retrovirus PERV-B6 env gene for envelope protein, genomic RNA.				
ACCESSION	AJ288591.1	GI:7414451			
VERSION	env gene; envelope protein.				
KEYWORDS	porcine endogenous retrovirus.				
SOURCE	porcine endogenous retrovirus.				
ORGANISM	Viruses; Retroviral viruses; Retroviridae; Mammalian type C retroviruses; 1 Mammalian type C virus group.				
REFERENCE	1 (bases 1 to 2002)				
AUTHORS	Bosch, S., Arnould, C., and Jestin, A.				
TITLE	Study of full-length porcine endogenous retrovirus genomes with envelope gene polymorphism in a specific-pathogen-free large white swine herd				
JOURNAL	J. Virol.	74 (18),	8575-8581	(2000)	
MEDLINE	20411432				
REFERENCE	2 (bases 1 to 2002)				
AUTHORS	Bosch, S.				
TITLE	Direct Submission				
JOURNAL	Submitted (28-MAR-2000)	Bosch S.,	Molecular Biology Department,		
FEATURES	AFSA-zoopole, zoopole les Croix B.P.53, Ploufragan 22440, FRANCE				
Source	Location/Qualifiers				
	1..2002				
	/organism="porcine endogenous retrovirus"				
	/specific_host="Sus scrofa"				
	/db_xref="taxon:61673"				
	/clone="PERV-B6"				
	/note="French Large White SPF-pig"				
	1..864				
	/gene="env"				
	1..864				
	/gene="env"				
	/note="truncated, not functional"				
	/codon_start=1				
	/product="envelope protein"				
	/protein_id="CAB86210.1"				
	/db_xref="GI:7414452"				
	/translation="MGPRSSRRLLPNNGGEPKRLRPLSFASIAFLTFNSDPOASSKRLIDSSDPHPLSLTWLIPDDIDVTNSTRGVAPRGTWPELHFCRLINPAVKSTPPNLVRSYGFPCPEETKEEYCGSGRSGFSCVSTNDGQWKPISLDPRVKSFPVNGSPKRYKVMELYKDKSCSPSDIDYLTISPFEKQKQENIKNMSGCIYFKKGGAAGSTLTIRLETGTGTEPPVAVGPNIVLAEGCPALIEPHNLPVPLISLRDITDPPDNSTTGISYQVNA"				
BASE COUNT	531 a	495 c	503 g	472 t	1 others
ORIGIN					
alignment_scores:	Quality: 3027.50	Length: 647			
	Ratio: 4.931	Gaps: 4			
Percent Similarity:	94.900	Percent Identity:	90.417		
Alignment_block:					
US-09-171-553B-6 x	PEN288591	..			
Align seg 1/1	to: PEN288591	from: 1	to: 2002		
1	MethSProthhrLeuSerArGargHisLeuProthArGlyGlyGluPr	17			
1	ATGGGTCCACGTCAGACCGGGCGCACCTCCGGAATGAGGGAGAGACC	50			
17	OLysArGLeuArGlieProLeuSerPhAlaSerlieAlaIrrppheLeuT	34			
51	GAAGAAGATTGAGAACCCCTTAAGCTTGCCCTCAATGCGCTGAGTCCCTTA	100			
34	hrLeuThrlleThrProGlnAlaSerSerlyArGleuIleAspSerSer	50			
101	CTTTTAACATGATGCCAGCCAGTACTAAGCGCTTAATGACAGCTCG	150			
51	AsnProHisArGProLeuSerLeuThrTrpLeuIlelleAspProAspTh	67			

151	GACCCCCAATAGACCCCTTATACCCCTTACGCGGTGATTAATGACCCCTGATAC	200
67	rclyValAlThrValAsnSerThrArglyValAlaProarglyThrTrpT	84
201	GGGTGTCACGTGTAATATACCTCAGAGGTGTGGCTCCAGAGCCACCTGGT	250
84	rProGluLeuHisPheCysLeuArqLeuIleAsnProAlaValLysSer	1000
251	GGCTGAACTGGATTTCGGCTCCGATTGATTAAACCCCGCTGTTAAAGC	3000
101	ThrProProAsnLeuValArgSerTyrGlyPheTyrCysCysProglyTh	117
301	ACACCTCCCAACCTAGTCCGAGTTAAGGGGTTCTATTGTGCCACAGCAC	3500
117	rgLyuysGluTyrTyrCysGlyYglYserGlyGluSerPheCysArgArqT	134
351	ACAGAAAGAGAAATTCGTGGGGGTTCCAGGAATCTCTGTGGGAGAT	4000
134	rPserCysValThrSerAsnAspGlyAspTrpTrpProIleSerLeu	150
401	GGAGTGGCGTACCTCCACAGATGAGACGTGAATGGCGATCTGCTC	450
151	GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrIly	167
451	CAGGACGGGGTAATAATCTCCTTTGTCAATCCGGCCCCGGGCAATGACAA	500
167	smetLysLeuTyrTyrAspLysAspLysSerCysSerProSerAspLeuAspT	184
501	AGTGATGGAACATAATAAGATAAGACGTCTCCCATCAGACTTGATT	550
184	YrLeuLysIleSerPheThrGlu.ArgLysThrGlyLysTyrSerLysV	200
551	ATCTTAAGAAATAGATTCTCACGAAAAAGAAACGGAATAATTCNAAG	600
200	alaSPlySTrPYrGluLeuGlyAsnSerPheLeu.LeuTyrGlyYglYgl	216
601	TGGATTAATGATGATGAGCTGGGAAATAGTTTTTAATAATATGAGAGGGG	650
216	yAlaGlySerThrLeuThrIleArgLeuArqIleGluThrGlyThrGluP	233
651	AGCAGGGTCCACTTACCTTACCTTGGATGAGATGAGACGGGAGCAGAAC	700
233	roP.roValAlaMetGlyProAspLysValLeuAlaGluIndlyProPro	249
701	CCCCGTGGGAGAGTGGGAGCCCAACATATGTAAGTGGAGAGGGGCCCGG	750
250	AlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuAr	266
751	GCCTGGAGGCCACCAACAACCTGGCGGTGCCAGTTAACCTGCTGG	800
266	gProAspIleThrGlnProProSerAsnSerThrThrGlyLeuIle.Pro	282
801	GGCTGAGATTAACACACGCCCTGACACAGTACCACTGGATTGATTTCCT	850
283	ThrAsnThrProArgAsnSerProGlyValProValLysThrGlyGlnAr	299
851	ACCAACACGGCTAGAAACATCCCCGGGTGTCTGTTAAGACAGCAGAG	900
299	gLeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAsp	316
901	ACCTCTCAGCCCTCATCCAGAGGAGCTTCCAAACCCATCAACTCCACCGAC	950
316	roAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTrp	332
951	CCGATGCCACTTCTTATTGTGGCTTTGTCTATCTCAGGGGCTCCCTTAT	1000
333	TyrGluGlyMetAlaLysGluArqLysPheAsnValThrLysGluHisAr	349
1001	TATGAGAGGAATGGCTAGAGAAAGGAATTCATATGAGCAAAAGAGCATAG	1050
349	gAsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValS	366
1051	AAATCAATATGATGGGGGTCCCGAAATATACCTCACCTCCACGAAAGTGT	11000

```
366  erGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeu 382
|||||
1101  CCGGGAAGGACATGCATAGAAAAGCTCCCATCCACCAACACCTT 1150
383  CysTyrSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeuVal 399
||| |||||
1151  TGCAATAGTACTATGTTTATAGCAGCGCTTACAAAATCATGATTTAGT 1200
399  LProGlyTyrAsnArgTyrPheAlaCysAsnThrGlyLeuThrProCysVal 416
|||||
1201  ACCTGGTATTAACATGATGGGATGCAATACAGGCTTAACCTCCCTGTG 1250
416  aLserThrSerValPheAsnGlnSerLysAspPheCysValMetValGln 432
|||||
1251  TTCTACACAGCTCTTACCAATCCAAAGATTCTG.GTCATGGTCCAG 1299
433  lLeValProArgValTyrTyrHisProGluGluValValLeuAspGluTyr 449
::: |||||
1300  CTCGTCCTCCCTGGTACACTACCATCCTGAGGACAGANWCATGCATGAGTA 1349
449  rAspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAla 466
|||||
1350  GGACTATCGGCTAACCCGCTCAAAAAGAGACCCGTCAGCCTAACCCCTAG 1399
466  lAValMetLeuGlyLeuGlyThrAlaValGlyThrGlyThrAla 482
|||||
1400  CCGTTATATCTAGGATTAGGGATGGCT...GGTGGGGACAGGACAGCT 1446
483  AlAlaLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuGlyLysLeuH1 499
|||||
1447  GCCCTGATCACAGACACAGCAGCTAGAAAAGACTGGTGAGCTACA 1496
499  sAlAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnL 516
|||||
1497  TCGCGCCATTACAGAGATCTCCAGCCTTAGAGGAATCTGTTAACAAAC 1546
516  euGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 532
|||||
1547  TAGAAGAAATCCCTGACTCT.TTGTCTGAGTGGTTCTACAGACTGGAGG 1595
533  GlyLeuAspLeuLeuPheLeuArgGluGlyLysLeuCysAlaAlaLeuLys 549
|||||
1596  GGATTAGATCTGCTGTTCTTAAAGAAAGTGGTTATGTGCAGCCTTAAG 1645
549  sGluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerM 566
::: |||||
1646  AGAAGAAATGTGCTTCTATGTGATCACTCAGAGCCATCAGAGACTCCA 1695
566  eTAsnLysLeuArgLysLysLeuGluArgArgArgGluArgGluAla 582
|||||
1696  TGAGCAACTTAGAAGAAAGTTAGAGAGACTCCAGAGGAGAGAGAGAGCT 1745
583  AspGlnGlyTyrPheGluGlyTyrPheAsnArgSerProTyrMetThrTh 599
|||||
1746  GACCAAGGGGTGTTGAAGAGATGTTCAACAGAGTCTCCTTGATGACCAC 1795
599  rLeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeuT 616
|||||
1796  CCTGCTCTCTGCTGATGGAGACCCCTAGTGTCTGCTCTGTTGCTTA 1845
616  hrValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArg 632
|||||
1846  CAGTTGGGCTTGGCTTAATTAATAGGTTATTGCTTTGTTAGAGAAATGA 1895
633  ValSerAlaValAlaGlnIleMetValLeuArgGlnGln 644
|||||
1896  GTGAGTGCAGTCAGATCATGTGTAAGACAGGC 1931
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2002, 02:53:17 ; Search time 10411.9 Seconds

(without alignments)
4683.669 Million cell updates/sec

Title: US-09-171-553B-9
Perfect score: 2956
Sequence: 1 tgcctttaggttagaac.....aaaaaaaaaaaaaaaaaa 2956

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenBank1:
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vl:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htg_hum:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2899.6	98.1	8750	14	PEN133818
2	2896.8	98.0	8196	6	A6552
3	2896.4	98.0	8209	6	A6553
4	2893.8	97.9	8763	14	PEN293657
5	2893.8	97.7	3482	6	AX002804
6	2889.4	97.7	3482	14	PEN17013
7	2839.6	96.1	7808	14	PEN133816
8	2811.6	95.1	8918	14	PEN133817
9	2004.4	67.8	8849	14	PEN133817
10	1970.8	66.7	1974	6	A6559
11	1941.6	65.7	1971	14	AF014162
12	1882.8	63.7	7333	4	AF038601
13	1882.8	63.7	7333	6	AR130474
14	1868.6	63.2	2000	14	PEN288592
15	1868.4	63.2	2000	14	PEN288588
16	1854	62.7	1968	14	PEN288589
17	1735.8	58.7	1996	14	PEN288590
18	1696.2	57.4	1999	14	PEN288587
19	1695.4	57.4	2002	14	PEN288591
20	1661.6	56.2	2008	14	PEN288586
21	1563.8	52.9	8918	14	PEN293656
22	1563.4	52.9	2462	6	AX002802
23	1552.2	52.5	2462	14	PERENV1
24	1552.2	52.5	4918	6	AX052637
25	1490.8	50.4	8132	4	AF038600
26	1489.2	50.4	8132	6	AR130475
27	1483.8	50.2	8132	4	AF038599
28	1433.8	48.5	7873	6	AX052638
29	1231.8	41.7	8060	6	AR130473
30	1053.2	35.6	1996	14	PEN288584
31	1009.8	34.2	1980	6	AX052633
32	1009.8	34.2	1980	14	AF130444
33	1007.4	34.1	2009	14	PEN288585
34	987.8	33.4	1840	4	AF296168
35	719	24.3	7362	6	AX052634
36	622.8	21.1	4402	6	AX052635
37	600.2	20.3	3320	6	A6551
38	545.2	18.4	1042	14	PERY17012
39	519.4	17.6	158455	2	AC079471
40	485.2	16.4	8379	14	GL060065
41	478.4	16.2	3025	4	AF147808
42	466.6	15.8	8431	4	AF151794
43	460.8	15.6	590	14	PEN298074
44	455	15.4	6312	6	A60211
45	455	15.4	6312	6	AR122287

ALIGNMENTS

RESULT 1
LOCUS PEN133818
DEFINITION Porcine endogenous retrovirus type C proviral gag, pol and env genes and LTR (class B, clone 43).
ACCESSION AJ133818
VERSION AJ133818.1 GI:6688949
KEYWORDS env gene; env protein; gag gene; gag protein; pol gene; pol protein.
SOURCE porcine endogenous retrovirus.
ORGANISM Porcine endogenous retrovirus.
VIRUSES; Retroviral viruses; Retroviridae; Mammalian type C retroviruses; 1 Mammalian type C virus group.
REFERENCE 1 (bases 1 to 8750)
Czudecna, F., Fischer, N., Bolter, K., Kurth, R. and Tonjes, R.R. Establishment and characterization of molecular clones of porcine endogenous retroviruses replicating on human cells
JOURNAL J. Virol. 74 (9), 4028-4038 (2000)
MEDLINE 20219394
REFERENCE 2 (bases 1 to 8750)

D	6516	AACGATGAGACTGGAAATGGCCGATCTCTCCAGAGCCGGTAAATTCCTTTGTC	6575
Q	841	aattccgccccgggcaagtacaaaatgataaataataaagctctccca	900
D	6576	AATTCCGGCCGGGCAAGTACAAAGTATGATAAATATTAAGATTAAGACTCTCCCA	6635
Q	901	tcagactatgatatacacaagaatagtttctacgaaagaagaagaataatcaca	960
D	6636	TCAGACTTATGATATCTAAAGATTAAGTTTCACTGAAAAGCAAAACGGAATATTCA	6695
Q	961	aagtgcataaagtatagctgctgggaaatagttttaaataatagcggggaagca	1020
D	6696	AAGTGGATTAATGATAGACTGGGAAATGTTTATTAATATGCGGGGAGCAGAGG	6755
Q	1021	tcacacttaacactcgccttaagatagagacgggaacaacaccccttgggaatgga	1080
D	6756	TCACCTTTAACATTCCTCTTAAGATAGAGACGGGACAAACCCCTTGAGCAGTGGGA	6815
Q	1081	ccccataaagctacgtgcagacagggccccggccctggagccacgcaatactgccc	1140
D	6816	CCCCATAAAGTACTGGCTGAACAAGGGCCCCCGCCCTGGAGCCACCCGATTAATTGGCG	6875
Q	1141	gtgcaccaataaactgcgtgcggcctgacataacacagccgcttagcaacagtaacc	1200
D	6876	GTGCCCAATTAACCTGCTGCGGCTGACATTAACACAGCGCCCTTAGCAACGGTACCCT	6935
Q	1201	ggatgtactcctacacacagcctagaagaactcccaagtgcttcctgttaagacagaa	1260
D	6936	GGATTGATTCCTTACCAACAGCGCTAGAAACTCCCCAGGTCTCTGTTAAGACAGGACG	6995
Q	1261	agactctcagctcaccacagggagcttccaaagccacacacacacacacacacacac	1320
D	6996	AGACTCTTCAAGTCTCATCCAGGGAGCTTTCCAAAGCCATCAATCCACGACCGCTGATGCC	7055
Q	1321	actctctctgttgcttctctctctcctcagggcctccttataatgagggatgagctaa	1380
D	7056	ACTTCTTCTGTTGGCTTGTGTCTTATCCCTCAAGGGCTCTTATTAATGAGGGATGGCTAAA	7115
Q	1381	gaaagaataatcaatgtgcacaaagacacagaaatcaatcaatcaatcaatcaatcaat	1440
D	7116	GAAAGAAATTCATGTGACCAAAAGACATAGAAATCAATGTACATGGGGGTCCCAAT	7175
Q	1441	aagcttaccctcactgaagttccgggaagggagacatgacatgaagaaagctccccaac	1500
D	7176	AAGCTTACCTGATGAAAGTTTCCGGAGGGACATGATGAGAAAGCTCCCCCATCC	7235
Q	1501	caccacaacacttgcctatagctgctgttataagacagccctcagaataatcagtatta	1560
D	7236	CACCAACAACCTTGTCTATAGTACTGTGTTATGACAGGCTTCAGAAATCAGTATTTA	7295
Q	1561	gtaccctgttatacaggtgtgagcatgcaatactggttaacccctgtgttccac	1620
D	7296	GTACCTGTGTTATACAGAGTGTGGCATGCAATACTGGTTAAACCCCTGTGTTCTCTCC	7355
Q	1621	tcagcttccaac	1680
D	7356	TCAGTCTTCAACCAATCAAAAGATTCTGTGTATGTGTCATGTCGCAAAATCGTCCCGAGTGTAC	7415
Q	1681	taccatccctgaaggaagtgctcctgataagataagactatcgatataacgcaacaaaga	1740
D	7416	TACCATCTCTGAGAAAGTGTCTCTGATGAATATGACTATGATGATTAACCAACCAAAAGA	7475
Q	1741	gaaccgtatcccttacccttagctgtatagctgagataggaagggccttggcgtgaga	1800
D	7476	GAACTCCGATCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC	7535
Q	1801	acagggacagctgcctgatacagacacacacacacacacacacacacacacacacacac	1860
D	7536	ACAGGAGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	7595
Q	1861	catcgccatgacagagaatctccgagcctlaaaggagctcgtttagcaactagaagag	1920
D	7596	CATCGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	7655

Q	1921	tccttactcttctgtctgaagtggtctacagaaacggagggatagatctgctgt	1980
D	7656	TCCCTGACTCTTCTTGTCTTAAGTGTCTTACGAAACCGAGGGATGATGATGATGATGAT	7715
Q	1981	ctaaggaagtggtgtatgtgacgcttaagaagaatgtgctctatagatgaac	2040
D	7716	CTAAGAGAGTGGGTATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	7775
Q	2041	tcagagacacataagagac	2100
D	7776	TCAGGAGCCATGAGACTTCATGAGCAAGCTTAAAGAAAGCTTAAAGAGGCTGAGAGG	7835
Q	2101	gaaagagagctgcagcaggggtgtgttgaagatgttcaacaggtctcctgtgaac	2160
D	7836	GAAAGAGAGCTGACAGGAGGTGTGTAAAGATGTTCAACAGGTCTCTGATGATGATGAT	7895
Q	2161	accctgtctctgtctgcaggggccccctagtcagtcgtctgtcttaacagttgg	2220
D	7896	ACCCTGCTTCTGCTCTGAGGGAGCCCTAGTATGCTGCTCTCTCTCTCTCTCTCTCTCT	7955
Q	2221	ccttgcttaataataggt	2280
D	7956	CCTTGCTTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	8015
Q	2281	atgtacttaggcaacagtaacagcccttgcagccacaggaagaaactgacactagct	2340
D	8016	ATGTACTTAGGCAACAGTACCAAGGCTTTCAGCAAGGAAAGTAACTGACTTACCTGATC	8075
Q	2341	tcacagcttaagatagaactataacagaacaggaagtgaggatgaagaatgaataa	2400
D	8076	TCCCACTTCTAAGATTAAGTAACTTAACAGCAAGAACTGGGAAATGAAGATGAAGAA	8135
Q	2401	tgcacactaacctcccaagacacaggaagtaataaagaagcttaaatgccccgaatt	2460
D	8136	TGCACCTTAACCTCCCAAGACCCAGGAAGTTAATTAAGAGCTTAAATGCCCCGAAAT	8195
Q	2461	acagaccctgcctgcagtaacatgaagtaagagctacacacttcttcttctcagggc	2520
D	8196	CCAGACCTGCTGAGGCTGCCAGTAATAGTAGAGGTACACTCTCTATGTTCCAGGGC	8255
Q	2521	ctgtactcctgcctgaatgaataacaggaatgaatgaatgaatgaatgaatgaatgaat	2580
D	8256	CTGTACTCTGCTGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG	8315
Q	2581	ctgtaaactgactgac	2640
D	8316	CTGTAAACTGACTGACACATAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	8375
Q	2641	tctcaactgcaatctgtcaactctgcaggaagacacacacacacacacacacacacac	2700
D	8376	TCTCAACTGCAATCTGTCACTCTGCCAGAGGCCACGAGATGGGAGCTCCGGAGCTA	8435
Q	2701	ttttaaatgattgttcaacaggaagcgtctcgtatattttaaataatgattgtccatg	2760
D	8436	TTTTAAATGATTGTGTCACAGGCGCGGCTCTGATTTTAAATTAATTAATTAATTAATTAAT	8495
Q	2761	gagcgcggtcctcgtatattttaaataatgattgttgaacacaggaagcttgggtgaa	2820
D	8496	GAGCGGGCTCTCGATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT	8555
Q	2821	ccccataaagctgtcccatctccgacacagggcccgagctccttaccctgtggtg	2880
D	8556	CCCCATAAAGCTGTCCGATTCGACATCCGAGCGCGGCGGAGTCTTACCCCTGCGTGGT	8615
Q	2881	tacgactgtggcccccagcgtctggaataaataatcctcttgcgtgttgaacaaa	2938
D	8616	TACGACTGTGGGCGCCAGCGCTTGGAATAAATCTCTGCTGTTGATGATCAAGA	8673

RESULT 2
A66552
LOCUS A66552 8196 bp DNA PAT 29-MAR-1999

DEFINITION Sequence 2 from Patent WO9740167.
ACCESSION A66552
VERSION A66552.1 GI:4538105
KEYWORDS
SOURCE porcine endogenous retrovirus.
ORGANISM porcine endogenous retrovirus
retroviruses; Retroviral viruses; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
REFERENCE
AUTHORS Galbraith, D.N., Haworth, C., Lees, G.M. and Smith, K.T.
TITLE PORCINE RETROVIRUS
JOURNAL Patent: WO 9740167-A 2 30-OCT-1997;
Q ONE BIOTECH LTD (GB)
FEATURES
Location/Qualifiers
1..8196
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
BASE COUNT 2165 a 2061 c 2146 g 1820 t 4 others
ORIGIN

Query Match 98.0%; Score 2896.8; DB 6; Length 8196;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 2937; Conservative 3; Mismatches 10; Indels 6; Gaps 3;

QY 1 tgctttaaagggttagaacaacccctgagacagtttggtgacccctatgaattgtct 60
DB 5246 tgctttaaagggttagaacaacccctgagacagtttggtgacccctatgaattgtct 5305
QY 61 acgggggagacccccccctggttgaagaattgtcttctgacatagctgagtgtgtctt 120
DB 5306 acgggggagacccccccctggttgaagaattgtcttctgacatagctgagtgtgtctt 5365
QY 121 cccagccttctgtctctagctcagagcgctcagagtggtgtaggcaagcgtgtagc 180
DB 5366 cccagccttctgtctctagctcagagcgctcagagtggtgtaggcaagcgtgtagc 5425
QY 181 agctccggagagctactcagagagagagagagagagagagagagagagagagag 240
DB 5426 agctccggagagagctactcagagagagagagagagagagagagagagagagagag 5482
QY 241 gagatcagctcagctcagagcagcagctgaggaagacacacacacacacacacac 300
DB 5483 gagatcagctcagctcagagcagcagctgaggaagacacacacacacacacacac 5542
QY 301 ctatctcgtactcttgac 360
DB 5543 ctatctcgtactcttgac 5602
QY 361 tccatgcatcccaagctgaagctgagcagcagcagcagcagcagcagcagcagcag 420
DB 5603 tccatgcatcccaagctgaagctgagcagcagcagcagcagcagcagcagcagcag 5662
QY 421 ctggaatcccttaagctcgtcctcagcagcagcagcagcagcagcagcagcagcag 480
DB 5663 ctggaatcccttaagctcgtcctcagcagcagcagcagcagcagcagcagcagcag 5722
QY 481 caggcagctgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 540
DB 5723 caggcagctgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5782
QY 541 tggctgattatgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 600
DB 5783 tggctgattatgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5842
QY 601 agaggcagctggtgagcagcagcagcagcagcagcagcagcagcagcagcagcag 660
DB 5843 agaggcagctggtgagcagcagcagcagcagcagcagcagcagcagcagcagcag 5902
QY 661 agcacacctcccaacagctagctcagctcagctcagctcagctcagctcagctcag 720
DB 5903 agcacacctcccaacagctagctcagctcagctcagctcagctcagctcagctcag 5962

QY 721 gagaatactgtggtgtctgtggaatcctctctgtagagagagcgtgcacctcc 780
DB 5963 gagaatactgtggtgtctgtggaatcctctctgtagagagagcgtgcacctcc 6022
QY 781 aagatgagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 840
DB 6023 aagatgagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 6082
QY 841 aatccggccggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 900
DB 6083 aatccggccggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 6142
QY 901 tcagactatgatactcagaagagagagagagagagagagagagagagagagagag 960
DB 6143 tcagactatgatactcagaagagagagagagagagagagagagagagagagagag 6200
QY 961 aagtgagatagatgagcagcagcagcagcagcagcagcagcagcagcagcagcag 1020
DB 6201 aagtgagatagatgagcagcagcagcagcagcagcagcagcagcagcagcagcag 6259
QY 1021 tccacttaacacatctgctcagtagagagagagagagagagagagagagagagagag 1080
DB 6260 tccacttaacacatctgctcagtagagagagagagagagagagagagagagagagag 6319
QY 1081 cccgataaagctagctgagcagcagcagcagcagcagcagcagcagcagcagcagcag 1140
DB 6320 cccgataaagctagctgagcagcagcagcagcagcagcagcagcagcagcagcagcag 6379
QY 1141 gtagcccaatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1200
DB 6380 gtagcccaatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 6439
QY 1201 gtagtattcctcacaacagcagcagcagcagcagcagcagcagcagcagcagcagcag 1260
DB 6440 gtagtattcctcacaacagcagcagcagcagcagcagcagcagcagcagcagcagcag 6499
QY 1261 agactctcagctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1320
DB 6500 agactctcagctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 6559
QY 1321 actctctcagctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1380
DB 6560 actctctcagctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 6619
QY 1381 gaaagaaatcagatgacacagcagcagcagcagcagcagcagcagcagcagcagcag 1440
DB 6620 gaaagaaatcagatgacacagcagcagcagcagcagcagcagcagcagcagcagcag 6679
QY 1441 aagcttaacccagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1500
DB 6680 aagcttaacccagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 6739
QY 1501 caccacaacactgtcagatagctgagctgagcagcagcagcagcagcagcagcagcag 1560
DB 6740 caccacaacactgtcagatagctgagctgagcagcagcagcagcagcagcagcagcag 6799
QY 1561 gtagcgtgtatatacagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1620
DB 6800 gtagcgtgtatatacagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 6859
QY 1621 tcagcttcaac 1680
DB 6860 tcagcttcaac 6919
QY 1681 taccactcag 1740
DB 6920 taccactcag 6979
QY 1741 gaaccgtatcccttaac 1800
DB 6980 gaaccgtatcccttaac 7039
QY 1801 acaggagacagctgcctgagcagcagcagcagcagcagcagcagcagcagcagcagcag 1860

Db 7040 ACAGGGACAGCTGCCATGATCACAGACCAAGACAGCTAGAGAAAGACTTGGTACCTA 7099
 Qy 1861 catggccatagacagaagatctccagccttaagaagctctgttagcaacctagaagag 1920
 Db 7100 CATGGCCATGACAGAGAGATCTCCGAGCCCTTAAGGAGTCTGTACCAACCTAAGAGAG 7159
 Qy 1921 tccctgactctctgtctgaagtggtcttcacagaaccggagggatagatctctgttt 1980
 Db 7160 TCCTTGACTTCTTGTCTGAAGTGTCTCTACAGACCGGAGGGATTAAGTTCGTCTTT 7219
 Qy 1981 ctaagaagaagtggtgtatctgcaagccttaagaagaatgctgtctctctctctac 2040
 Db 7220 CTAAGAGAGAGTGGGTATGTGACGCTTAAGAAAGAAATGTTGCTTCTATGTAGATCAC 7279
 Qy 2041 tcaagaagcctcaagaagctcatgaagaacttagaagaagaatgtagaagagctcgaag 2100
 Db 7280 TCAGGACCCATCAGAGACTCCATGACAACTTAGAAAAAGTTAGAGAGGCGTCCAGAG 7339
 Qy 2101 gaaagaagaagctgacacaggggtgttgaagaagtgttcaacaggtcctctgtgagac 2160
 Db 7340 GAAAGAGAGGCTGACAGAGGGGTGTTGAAGGATGTTCAACAGGTCTCTTGGATGACC 7399
 Qy 2161 accctgcttctgctctgacggggcccttagtagtccctgctcctgtacttaagttgag 2220
 Db 7400 ACCCTGCTTCTGCTGACGGGGCCCTAGTAGTCTCTGCTGCTTACTTACAGTTGGG 7459
 Qy 2221 ccttgcttaataagtggtgtgtgcttctgttagagaagaatgtagtccagatc 2280
 Db 7460 CTTGCTTAATTAATTAAGTGTGTTGCTTGTAGAAAGAAAGTGAAGTGAAGTCCAGATC 7519
 Qy 2281 atgtacttaggcaacagtaaccaaggtctctgacccaagaagaactgacctagacct 2340
 Db 7520 ATGTTACTTAGGACACAGTACCAAGGCTTCTGAGCCAGAGAAACTGACTCTTGCCCT 7579
 Qy 2341 tcccaagcttaagaattagaacttaacaagaagaagtggggaaatgaagaatgaaga 2400
 Db 7580 TCCAGTTCTTAAGATTAGAACTATTAAACAAGACAGAAAGTGGGAATGAAGAGATGAAAA 7639
 Qy 2401 tgaaccttaacccccaagaacccaagaagaatgaataaagaagctcctaagccccgaat 2460
 Db 7640 TGCACTTAACCTCCAGAACCCAGAAAGTAAATAAAAAGCTTAATATGCCCCCAAT 7699
 Qy 2461 acaagacctgctgctccagtaataagtagaaggtcaacactctatgttccagggc 2520
 Db 7700 MCAGACCTCTGCTGGCTGCCGTAATAGTAGAAGGTACACTTCTTATTTGTTCCAGGGC 7759
 Qy 2521 cgtctatccctgacctaaagtaagaatacaagaaatgagttagtaactcgtctatctgagt 2580
 Db 7760 CTGCTATCCTGGCCTAAGTAGAATTAACAGGAATGAGTTGACTAATCGCTTATCTGATT 7819
 Qy 2581 ctgtlaaactgactgagcacatagaagaatgattcaacattgacagccctagtgacct 2640
 Db 7820 CTGTAAAAAGACTGGACCATAGAAAGAAATGATTACATATGACAGCCCTATGTGACCTA 7879
 Qy 2641 tctcaactgcaactgctcaactgctccagaagaagcccaagagatgagagactccgagact 2700
 Db 7880 TCTCACTGCAATCTGTCACTCTGCCAGAGAGCCCAAGATGCGGACTCCGGAGCTA 7939
 Qy 2701 ttttaaaatgagtgttccacagagcgaggctctcgtata ttttaaaatgagtgttccatg 2760
 Db 7940 TTTTAAAAATGATGTGTCACAGGAGCGGCTCTCGATATTTTAAATGATTGGTCCATG 7999
 Qy 2761 gaagcgagggtctcgtatatttttaaaatgattgtgtgagcagcagagcttggtgtga 2820
 Db 8000 GAGCGGGGCTCTCGATATTTTAAATGATGGTTGTGACGACAGGCTTGTCTGTGA 8059
 Qy 2821 ccccataaagctgtcccgatcccgactcgagcgagtcgagtccttaacccctgagtgag 2880
 Db 8060 CCCCATAAAGCTGTCCGATTCGACATCGGGCGGACGATCTCTACCCCTGCGTGTG 8119
 Qy 2881 taagactgtgggccccagcgagcttgaaataaaatcctctgtgctgttgatcacaanaa 2940

Db 8120 TACGACTGTGGGCCCGACGCGCTTGATTAATAATCTCTTGTGATCAAAAAA 8179
 Qy 2941 aaaaaaaaaaaaaaaaaa 2956
 Db 8180 AAAAAAAAAAAAAAAAAA 8195

RESULT 3
 A66553
 LOCUS A66553 8209 bp DNA PAT 29-MAR-1999
 DEFINITION Sequence 3 from Patent WO9740167.
 ACCESSION A66553
 VERSION A66553.1 GI:4538106
 KEYWORDS
 SOURCE
 ORGANISM
 porcine endogenous retrovirus.
 porcine endogenous retrovirus
 Viruses; Retroid viruses; Retroviridae; Mammalian type C
 retroviruses; 1-Mammalian type C virus group.
 (bases 1 to 8209)
 Galbraith, D.N., Haworth, C., Lees, G.M. and Smith, K.T.
 Patent: WO 9740167-A 3 30-OCT-1997;
 Q ONE BIOTECH LTD (GB)
 Location/Qualifiers
 1..8209

FEATURES
 source
 1..8209
 /organism="porcine endogenous retrovirus"
 /db_xref="taxon:61673"
 588..2162
 /codon_start=1
 /product="VIRAL CORE PROTEIN"
 /protein_id="CAB39347.1"
 /db_xref="GI:4538107"
 /translation="MGQIVTTPSLTLDHTVEVSRHAINLSVQKKPWPOTFCASNP
 TFDGWSBEGTFENSEITLIAVAIIIFGTGPGSHPOEPIITWQDLADDPWNPWLN
 KPRKGPRIILALGKRNKSHAKVPPSSSYLRDGGADLAGTPTCSPTPLSTGCCG
 TSAPGAPVAVPPAGTGRSRGATPERIDETALILPLTYGPMGGOLOLOLVYFSS
 ADLNKMTNHPFSEDPQRLGLVLSMFESHQPWDDCOQLOLPLTTEBERILLFA
 RKNVAGAGRTQLOLNETDMQFPPLTRGDMQNTAEGRESLKIYQALVAGLRGASRP
 TNLAKRVWQGNPPSPVLERLMEAFRRPTTDPDISEAKASVALAFIGOSALDJR
 KILQRLBGLQAEILRDVLREAEKVYRRETEERKEREERERERRORRQKNT
 KILAAVEGKSSREERDFRIRSGPROSGNLGNRTPLDKQCAVCKEKGWANNCPK
 KGNKPKVLALEEDKD"
 <2163..5747
 /codon_start=1
 /product="REVERSE TRANSCRIPTASE"
 /protein_id="CAB39348.1"
 /db_xref="GI:4538108"
 /translation="GRGSDPLPPRYTLKVEGQVPEFLVDTGAEHSVLLOPLKLE
 KSNWAGTGGROYPMTRRTVDLVGVRVTHSPVLPCEVPILGRDLIRKMAOISF
 BOGPEVSNVKKPTIVLTLOLDDEPRYLXSVQYKRPDOIOISLBFPDPAKETAAGMGA
 KOVPPVYIOLKASATPVSVROYPLSREARBEIMHVORLIQOGLIVVQSPWNPPLP
 VRKPTNDYRVPVDLRVKNRVODIHPVPPNXLISALPERRWYVLDLKAFFCL
 RLHPTSOPLFAFERMDGTGTGOLITRLPQGNKSPITFEDALHRLDLANFRIOHQ
 VTLTQYVDLILAGATKODCLEGRKALIELSDLGYSASAKAOICREYTVLGYSLR
 GGORMLTEARKKIVVOIIPAPTKAVEPELTGTAGFCFLMIPGKATLAAPLYPLTKKG
 GFSAPRHKARFADIRKALISAPLAPDVTKEPTLYVDRKRGARVLTOTGPWR
 PVATLSKLDPVASGWPVCLKALIAVALIVKADKLTIGONTIVIAHALENTIROP
 DRWMTNKRMTYOSLLITERTEAPRALNPATLLPEITDEPVIHDOHLLIEGTGR
 KDLDIPLTGEVLVMTFDGSSYYVEGRMAAAYVDGRTIMASLSEPTSAOAEIM
 ALTDALRLAEKSNINITYDSYAPATAGVIGATIKQGLITLSAGRELTKNEEITSLIE
 ALHLPLRLAIIHCGHOKANDLISRGNDMAVRYKQAOAVNLPLITTEPKAPBEP
 VYLEDMEIKIKIOFSEPTGCTCTGSKTILPKBELFVVOOHRTHLTHGTHLOOL
 VRTSPYHVLRLPGVADSVKHCPCOULVNNPNSRIPGKRLRSGRPAHWEVDTEK
 PAKYGNKYLVLVEFTESGWAEVPTKRETSVAKRLLEIPFPFGIPKIVGSDNGFA
 FVAQVSOGLAKILGIDMKLHCAYRPOSSGOVERNNRTIKETLRLTETGNIINMAEL
 PVLVLRNTPGGGILPYELLVGGPPLAEIIFAHSAVDLLSOLPLFSRLALEMWO
 RAMKOLREAYSGDLOVPHREFOYDSSVYVRRHNGNLSETWKGGVLYLTLTPAAVVE
 GIPLWIASHVKKAPPPDSQMRBKEKTENPLKLRHLRLVYPSNNNSGO"
 5620..7590
 /codon_start=1
 /product="VIRAL ENVELOPE PROTEIN"
 /protein_id="CAB39349.1"
 /db_xref="GI:4538109"

CDS

/translation="MNPILSRRLHPTRGCEPKRLRPLSPASIAWPLTLTTPPOASK
RLIDSNPRLPLSLWLLIDPDTGVNSTRGVNARGTWMLPLHLCLRLINAVSTP
PNLVRSTGCEPCEPTEKEXCGSGSCFRMSCTGNTMDKMLSLDIDRKFSEVN
SGBGKXMMKILKDKSCSPSDIDYLKISPERKTKYSVYDKWELGNFLLYGGGAG
STLTLRLIETGTEPPVAMPKPVLAEGPPALPPPHNLPVQULSLRPDITOPNS
TGLIETNPNSPGVPVATGQRLSLLOGAFOALINSDPDATSSCWCLJSGPYE
GNKERRFWNTKEHRNOCITWGSRNKLTLLEVSCKICIGKAPSHOHCYSTVVEOA
SEMOYLVPYCNMWMACNIGLPCVSTVSNOSKDCVMQIVPRYVHEEVLEEDY
YRNRPKREPVLITLAVMLGTAVGFTALITGPOLKEGELHAAETEDLRA
LKEVSNLLESLTSEVYLONRGIDILFLREGICLALKEGCEFYDHGATIDSM
NKLRLKLEPRRRRRREADQCFEGCFMRSRMTTSLALGPLVLILLTVPCILNR
FVAFVERYSANOYWLKROOIGLSQETDL"

BASE COUNT 2168 a 2064 c 2154 g 1823 t
ORIGIN

Query Match 98.0%; Score 2896.4; DB 6; Length 8209;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 2939; Conservative 0; Mismatches 11; Indels 6; Gaps 3;

QY 1 tgcctttaaggttaagaaacaccccttgacagtttggtctgacccctlaaalgctc 60
|||||
Db 5260 TGCTTTTAAAGGTAGAGAACACCCCTGACAGTTTGGGCTGACCCCTATGATGCTCT 5319
QY 61 acgggggagaccccccgttgatagaattgcttcgtacaagtgctgtagtgctt 120
|||||
Db 5320 ACCGGGGAGACCCCCCTTGCGAGAAATTCCTTGGACATAGCTGTGATGCTCTCTT 5379
QY 121 cccagctctgtctcctcagctcaagcgctgtagtggtgaggaacagcggtgaag 180
|||||
Db 5380 CCCAGCTTGTCTCTAGGCTCAAGGCGCTCGAGTGGTGAAGGACGACGCTGGAAGC 5439
QY 181 agctcggagagcctactcagagagagagacttgcaatgccatccgtccaaagt 240
|||||
Db 5440 AGCTCCGGAGGCGCTACTCAGAGAG--GAGACTTGCAAGTCCACATCGCTTCAAGTTG 5496
QY 241 gagatlaagctatgttagagcgacacgctgacagaaaccccgagctcgttgaaagag 300
|||||
Db 5497 GAGATTAGTCTATGTTAGAGCGCACCTGACAGAAACCTCGAGACTCGGTGAAAGGAG 5556
QY 301 ctatcctgcacttttagacacacacacagcgctgtgaagtcgaaggaatcccaactg 360
|||||
Db 5557 CTATCTCGTACCTTTTGACACACACACAGCGCTGTGAAGTCGAAGGATCCCACTGGA 5616
QY 361 tcaatgcatcccaagcttaagctgagcgacccctccgactcgggtgagagacggaaga 420
|||||
Db 5617 TCATGTGATCCACAGTTAACCGCGCGCACCTCCGACTCGGGGTGAGAGCGCAAGAA 5676
QY 421 ctgaagaatcccaatgaactcgcctcactcgtctgtcttacttaacaaataaactcc 480
|||||
Db 5677 CTGAGAAATCCCTTAAAGCTTCGCTTCATCGCTGGTCTGCTTACTTAACAAATTAAC 5736
QY 481 caagccagtaglaaacgcttatagacagctcgaaaccccatagacattatccctac 540
|||||
Db 5737 CAGGCCAGTAGTAAGCGCTTATAGACAGCTCGAACCCCATAGACTTATTCCTTAC 5796
QY 541 tgcgtgatattgacctgatacggtgtcaactgtataaagaactcgaagtgctgctc 600
|||||
Db 5797 TGGCTGATTTATGACCTGTATAGCGGTGTCACGTAAATGACACTGAGAGTGTGTCCT 5856
QY 601 agagacacctggtgactgaactgcatcttcgctcgaattgattaaaccccggttaaa 660
|||||
Db 5857 AGAGGACCTGGTGGCTGAGACTGCAATTCGCTCCGATGATTAACCCCGTGTAA 5916
QY 661 agacacactcccaactagtcgtagttatgggttctatgtcgcacagcaagaaga 720
|||||
Db 5917 AGCAGACCTCCCAACCTAGTCGTAATGAGTTCATGTGCTGCCAGGCAAGAGAAA 5976
QY 721 gagaataactggtggggttctggggaatcctctgtaagaagatgagctgctcaactcc 780
|||||
Db 5977 GAGAAATACGTGGGGGTTCTGGGGAATCCTCTGTAGAGAGATGAGAGCTGCGTCACTCC 6036
QY 781 aacgatgagacggaatgagcgcgcatctctccaggaacgggtgaaatctccttgctc 840

Db 6037 AACGATGAGAGATGGAATAGCCGATCTCTCCAGGACCCGGGTAAATTTCTCCTTGTG 6096
QY 841 aattcggcgccgggaagtacaaatgataaactataaagataaagatcgtcccca 900
|||||
Db 6097 AATTCGGCGCCGGGCAAGTACAAATATGATGAATATATTAAGATTAAGCTGCTCCCA 6156
QY 901 tcaagcttaagataactaaagaataagttcactgaaagaaagaaacagaaataatcaa 960
|||||
Db 6157 TCAGACTTATGATTAATTAAGATTAAGTTCACCTG--AAAGGAAAACAGAAAATTTCAA 6214
QY 961 aagtgataaactgataagctgagctggggaatgcttttaataatgctgggggagcagg 1020
|||||
Db 6215 AAGTGATTAATGATATGATGAGCTGGGGAATAGTTTTT-ATTATATGCGGGGAGCAAGG 6273
QY 1021 tccacttaacatctgccttaagataagagcgggagacaaacccctgttgcaatgga 1080
|||||
Db 6274 TCCACTTTTAACTATTCGCTTATAGATTAAGACGGGAGACAGAACCCCTGTGCAATGGGA 6333
QY 1081 cccgataaagtaactgctgaacaggggccccggtcctgagccacacgataacttgccg 1140
|||||
Db 6334 CCCGATAAAGTACTGTGTAACAGAGGCCCCCGGCTGAGGCCACCGCATTAATTTGCGG 6393
QY 1141 gtgccccaatlaactcgtcgtgagctgacataaacaagcgctagaaagaatcaact 1200
|||||
Db 6394 GTGCCCAATTAACCTCGTGGGCTGACATTAACACAGCGCGCTTAGCAACAGTACACT 6453
QY 1201 ggaattgattcctacacaacgctgaaactcccaagtgcttcctgtaagaaggaag 1260
|||||
Db 6454 GGATTGATTCCTACCAACAGCGCTGAAACTCCAGGATGTTCTGTTAAGACAGAGACAG 6513
QY 1261 agactcctcagctcatalccaggaagcttccaagccatcaactccacgacactgagc 1320
|||||
Db 6514 AGACTCTTCAGTCTATCCAGGAGACTTCCAGGCAATCACTCCAGGACCTGATGCTCC 6573
QY 1321 actcttcctgttggctgtgtcctcctcagggcctctttaaagggagatgctaaa 1380
|||||
Db 6574 ACTTCTCTGTTGGCTTTGTCTATCTCAGGCTCTCTTATATGAGGAGATGGCTTAA 6633
QY 1381 gaaagaaatlaaactgtaacaaagagacataaactcaatgacatgaggggtcccgaaat 1440
|||||
Db 6634 GAAAGAAATTAATAGTGAACCAAGAGCATAGAAATCAATGATGGGGGCCGGAAT 6693
QY 1441 aagcttaacctactggaagttccgggaaaggagacatgataagaaagatcccccaccc 1500
|||||
Db 6694 AAGCTTACCTCTACATGGAAGTTTCCGGGAAGGGAATCATAGGAAGAGCTCCCACTCC 6753
QY 1501 caacaacacctgtgctatagtaactgtgttataagagagcctcagaataatcatatta 1560
|||||
Db 6754 CACCAACACTTGTGCTATAGTACTGTGTTATGAGCAGGCTCAGAAAATCAGATTTA 6813
QY 1561 gtaactggtataaaggtggtgagcaltgcaataactggtttaaaccctgtgttccacc 1620
|||||
Db 6814 GTACCTGGTTATTAACAGGAGTGGGATGCAATACTGGGTTAAACCCCTGTGTTCCACC 6873
QY 1621 tcaagcttaacaaatcaaaagattgtgttcaatggtccaatctgtcccgagtgtaac 1680
|||||
Db 6874 TCAGTCTTCAACCAATCCAAAGATTTGTGTGCAATGAGTCCAAATGTCCTCCGAGTAC 6933
QY 1681 taccaactgaaggaagtgctccttgaataatgactcgtgataaacgcgaacaaaga 1740
|||||
Db 6934 TACCATCTCGAGGAAGTGGTCTTATGATATGACTATCGGATTAACCGACCAAAAAGA 6993
QY 1741 gaaccgtaactccttaacctagctgtaaigtctcgatataaggaagcgctgtgctgta 1800
|||||
Db 6994 GAACCCGATATCCCTTACCTAGCTGTATGTCGGGATTAAGGAGAGCGCGCTGGGATAGA 7053
QY 1801 acaggaacagctgcccctgataacagaccacaagcagctagaagaagaagctgtgtagcta 1860
|||||
Db 7054 ACAGGAGACAGCTGCCCTATATACAGACACAGCAGCTAGAGAAAGGACTTGAGAGCTTA 7113
QY 1861 catgagcaltgacagaagatctcagagccttaagaggtctgttaagaacctagaagag 1920
|||||

Db 7114 CATGGGCGCATGACAGAAATCTCCGAGCCTTAAAGAGTCTGTAGCACCTTAGAAGAG 7173
 QY 1921 tccctgactctctgtcgtgaagtgtctctacagaccggaaggtatgatctgttt 1980
 Db 7114 TCCCTGACTCTCTGTCTGAAGTGGTTTACAGAACCGGAGGGATTAGATCTCTGT 7233
 QY 1981 ctaagaagaagtgtgttatgtgcagccttaaaagaatgtgtctctatgtatgacac 2040
 Db 7234 CTAAAGAGAGGTGGTTATGTGCAGCCTTAAAGAAAGAAATGTTCTCTATGTAGATCAC 7293
 QY 2041 tcaagagccatcacagagacttcaatgaagaacttagaaaaagttagagagcgctcgaaag 2100
 Db 7294 TCAGAGCGCATACAGAGACTCCATACAGAGCTTAGAAAAAGTTAGAGAGCGCGTAGAG 7353
 QY 2101 gaaagagagcgtgaccaggggtgtgttgaagagatgtgtcaacaagcttccctgtatgac 2160
 Db 7354 GAAAGAGAGCGGTGACCAAGGGGTGTTGAAGAGATGTTCACAGAGTCTCTTGGATACC 7413
 QY 2161 accctgtctctgtcgtcgaagggccctagtagtctgtcctgtacttaacagtgtg 2220
 Db 7414 ACCCTGCTTCTGCTGACGCGGCGCCCTAGTAGTCTGCTGCTTACTTACAGTTGGG 7473
 QY 2221 ccttgcttaataatagttgtgtgcttgttagaagaagatgtgagtcagtcagatc 2280
 Db 7474 CTTTGCTTAATTATAGATTGTGTGCTTTGTTAGAGACAGAGTGCAGTGCACAGATC 7533
 QY 2281 atgttacttaggcacacagtagtaccagagcctctgagcgaagagaactgagccttagcct 2340
 Db 7534 ATGCTACTTAGGCAACAGTACAGAGGCTTCTGAGCCCAAGAGAACTGACTCTTAGCT 7593
 QY 2341 tccagttcttaagaattgaacttlaacaagaagaagtgtgggaatgtgaaagtgtgaaa 2400
 Db 7594 TCCAGTTCTTAAGATTAGACTATTACAGACAGAGAGTGGGAATGAAAGAGTAGTAA 7553
 QY 2401 tgcacactaaccccccagaagccagaagaagttaataaaaaagccttaaatgtccccgaatt 2460
 Db 7654 TGCACACTTAACCTCCCGAGAACCCAGGAAGTTAATAAAAAAGCTCTAATGCCCCGAGATT 7713
 QY 2461 acagagaccctgtgctgcagtaaatagtagaaggtgcacacttccattgttccagggc 2520
 Db 7714 ACAGACCCCTGCTGCTGCCAGTAAATAGTAGAAGTGCACACTTCTATTTGCCAGGCG 7773
 QY 2521 ctgtatccctggcctaagttagaataacaggaatgtgagtaactatogcttactgtgatt 2580
 Db 7774 CTGTATCTCTGGGCTTAAGTAGTAACAGGAAATGAGTTGACTAATCGCTTATCTGAGATT 7833
 QY 2581 ctgtaaaactgactgtgacacatagagaatgttatcacatgtgacagccttagtgacct 2640
 Db 7834 CTGTAAACTGACTGGCACCATAGAAAGAAATTGATTACACTTTGACAGCCCTTAGTGACCTA 7893
 QY 2641 tctcaactgtcaactgttcaactgtcccaagagcccaagcagatgtgagacctccgagact 2700
 Db 7894 TCTCAACTGTCAATCTGTCACTTGCCCGAGSAGCCCAAGAGATGGCGAGCTCCGGAGCTA 7953
 QY 2701 ttttaaaatgattgtgtcccaagagcgccggccttcgatattttaaaatgtgtgtccatg 2760
 Db 7954 TTTTAAATGATTGTGTCCAGGAGCGCGGCTCTCGATATTTTAAATGATTGTGTCCATG 8013
 QY 2761 gaagcgaggtctcgaattatttaaatgattgtgtgtgacagagaggtttgtgtgaa 2820
 Db 8014 GAGCGGCGGCTCTCGATATTTTAAATGATTGTGTGTGACGACAGGCTTTGTGTGTGA 8073
 QY 2821 ccccaataaagctgtcccgatctccgactcggcgaggtcctctacccctgtcggtgtg 2880
 Db 8074 CCCCATAAAGAGTGTCCCGATCTCCGACTCGGCGGCGGAGTCTTACCCCTGCGGTGGTG 8133
 QY 2881 tagactgtgtggcccgagcggtgtgtgataaaatccctctgtcggtgtgtgcatcaaaaa 2940
 Db 8134 TAGACTGTGTGGCCCGAGCGCGCTTGGAATAAAAATCCTTGTGTGTGACATAAAAA 8193
 QY 2941 aaaaaaaaaaaaaa 2956
 Db 8194 AAAAAAAAAAAAAA 8209

RESULT 4
 LOCUS PEN293657
 DEFINITION Porcine endogenous type C retrovirus proviral gag gene, pol gene and env gene, class B, clone 213.
 ACCESSION AJ293657
 VERSION AJ293657.1 GI:14275841
 KEYWORDS env gene; envelope; gag gene; group specific antigen; pol gene; polymerase; polypotein.
 SOURCE porcine endogenous type C retrovirus.
 ORGANISM Porcine endogenous type C retrovirus.
 REFERENCE 1 (bases 1 to 8763)
 Krach, U., Fischer, N., Cauderna, F. and Tonjes, R.R. Comparison of replication-competent molecular clones of porcine endogenous retrovirus class a and class b derived from pig and human cells
 J. Virol. 75 (12), 5465-5472 (2001)
 JOURNAL MEDLINE 21256017
 REFERENCE 2 (bases 1 to 8763)
 Tonjes, R.R. Direct Submission
 JOURNAL Submitted (31-JUL-2000) Toenjes R.R., Medical Biotechnology, Paul-Ehrlich-Institut, Paul-Ehrlich-Str. 51-95, Langen, 63225, GERMANY
 FEATURES
 source
 location/Qualifiers
 1..8763
 /organism="porcine endogenous type C retrovirus"
 /proliferal
 /specific_host="Sus scrofa"
 /db_xref="taxon:70540"
 /clone="213"
 /cell_line="PK15"
 /note="class B"
 1..630
 467
 /note="cap site"
 633..650
 1077..2651
 /gene="gag"
 1077..2651
 /gene="gag"
 /codon_start=1
 /product="group specific antigen"
 /protein_id="CAC39625.1"
 /db_xref="GI:14275842"
 /translation="MGOTVTPPLSLTLDHMTVRSRAHNLISVQKRGPMOTFCASEMP
 TFDVMPSEGFENSEIILAVKALIFGPGSHPODEYILTWDLAEPPPMVPMIN
 KPRGPRPILALGKRNKHAKEVPSPGIYPEIEPPTPEPPVPPPAQAVNG
 PSAPGAYVEGPAAGRSRGATPERTDELAIIPLTKTGPMPGGOLQPLQWPSS
 ADLNMKTNPSPESDPORLGLIVESLMFSHPQWDDCOQLDQLFTTERERILLER
 RKNVPGADGRTOLQNEIDMGFLTPRGWYNTAEGRSLKTYROALVAGLRGASRRP
 TNLAKREVMOGPMPEPVSVELERLMEAFRRFTPDPTSEAKQASVALAFGOSLDIR
 KILORLEGLQAELELDLVREAEKYYRRETEEEKOEKREERERERRRROBKNT
 KILAAVEGKSSRRERDPFRIRSGPQSGNLGRRTLDMDQCAVCKEKHMARNCX
 KNGKPRVLALEPKD"
 2652..6239
 /gene="pol"
 /db_xref="GI:14275843"
 /translation="GRGSDPLPEPPVTLKVEGOPYEFLVDTGAHSVLLIOPGRIKE
 KSMVKGATGQROYPTTTRTVDLAVGRVTHSLVPECPVLLGLRLTKGAQISF
 BOGRPEVSNKPLITVLTLDDEYRLYSQVAPDDIOEWLEPQVAAVEAGMGLA
 KQVPPVIOIKASATPVSRQYPLSRBARBGIVPHVORLIIOGILVPOVSPMWTPLLP

VRKPGTNDYRPOVDIREVNRKRODHPYTPNPYNLLSALPERNNYTVJDLDAFECFL
 RLHP10PLFAEWRDPTGRTGQLTWLTLPOGFKNSTPFEALRDLRANR10HPQ
 VTLLOVVDLLLAGATKODLETKALLLESLDGRASAKAOICRRTVLYGSLR
 GGGRLTEARKRTVVOIPAPTAKEVRELTGAFRLMIPGATLAAYLTLETKRG
 ESMADENHOKAFDAIKKALLSAPALALPVPTPELYVERGVAGVLTOLYEMBR
 PVAVLSKTLDPVAGSPVCLKAIAVAIILKADKLTGONITVIAAPHLNEMVOPR
 DRMTNARKTHYOSLLTERVTFAPPAALNPATLPEETDEVYTHDCHLLIEETGVK
 KDILIDPLTGEVLTWFTDSSYVVEGKAGAAVNDGTITWASSLPESTSAQKAEIM
 ALQALRLAEGKSINITYDSRYAFATAHVHGA1YKORGLITAGREIKKKEI1SLE
 ALHLPKRLA1IHCPGHOKAKDLISRGNOADRYAKOAAVNDLPIETPKAEPERQ
 YLLEDMOELKIDQSETEGTCYTSYKEILPHKEGLYQOIHRLITGLTKHLQOL
 VRTSPYHVLRLBEGVADSVYKHCPCOLVANDSRIPGKRLSGSHGAWEDPTEFK
 PAVGKMYLLVYDPTSGWVEAPYPRKESYVAKKILEIEIPRIGIPYVISDNGPA
 PVAQVSGIATKILGIDMKLHCATRPOSSQOVERMNTIETILTKLTTEGINDWIAL
 PVLFRVKNTPGQFGLTPELYELLYGPPPLVEIASVHASADVLISQPLFSKALEWVRQ
 RAWKOLREAYSGEGDLOVPHRFQVDSVVRHHAAGNLTPYKSGVLLTTPYAVKY
 EGISTWISHSHVLAAPPDPSGWRKAEKTEMLKRLRLHVLVPSNNNSPQ"

6112. .8085
 /gene="env"
 6112. .8085
 /gene="env"
 /codon_start=1
 /product="envelope"
 /protein_id="CAC39627.1"
 /db_xref="GI:14275844"
 /translation="MPTLSWRHLPTRGGEPRKRLRPLSFASIANFVLTITTPQSSK
 RLIDSSNPRLPUSLITWLTIDPDIGVNVSTKQVAPRGTPWPELHFLRLINAVASTP
 PNLVRYGFCYCCGTEKERKCYSGGSGESFCRWSCVTNSNGDKRWP1SLQDRKSEFN
 SGPGKRVKMLYKDKSCSPSDLYLKI1SPTEKQENIOKWTNGMSWIVPKYGGGA
 GSTLLIRLLETGTEPPVAGPDLVLAEGPAPLEPHMLPVOLPLSLDITTOPXS
 STGLILPTNPRNSPGVPVTKGQRLSLIOGAFOALNSTDPATSCMTLSGPPY
 EGMKKEKFNVTREHSGNCTGMSRNLKLTVEYSGCTGKAGKPSIOHLKENVYVEO
 ASENOYLVCVNNWMAKNTGLIPCVSTSVFNSKDPVCVVO1VPRYVHPEVYEDLX
 DYKINRKPBPVSLTAVMLGLTAVGVTGTPALITGQJLEKGIHGLHAAETDLR
 ALSEVSNLLESLTSLSEVQLNRRGLDILFLREGCLALKECCFYVDHGAIRDS
 MSKIRLERLRERREADOGWEGEFMNRPMWTTLISALTGLPVLVLLITLTVGPOLIN
 RFVAFVRENVSAQIOWMLROXYGGLISQETDL"

8133. .8763
 LNR
 polyA_signal 8655. .8660
 BASE COUNT 2301 a 2200 c 2287 g 1974 t 1 others
 ORIGIN

Query Match 97.9%; Score 2893.8; DB 14; Length 8763;
 Best Local Similarity 99.0%; Pred. No. 0;
 Matches 2910; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 1 tgctttaggttagaacaaccccttgacagatttggtctgacccctatgaattgctc 60
 |||
 Db 5749 TGCCTTTTAGGGTTAGGACACCCCTTGACAGTTTGCGCTGACCCCTATGATTCGCTC 5808
 QY 61 acgggggaaccccccgcttgtagaacttgctctcgtacataagtgctgactgctc 120
 |||
 Db 5809 ACGGGGGACCCCCCTTGATGAGAAATTCCTCTGACATCTGATGCTGCTCTCTT 5868
 QY 121 cccaagcttgcttcctgaagctcaagcgcttgagtggtggaagcaacggcgctgaagc 180
 |||
 Db 5869 CCCAGCCTTGTCTCTAGGCTCAAGCGCTCGAGTGGGACCAAGCGCTGGAGGC 5928
 QY 181 agctccggagagctactcagagagagagacttgcaagttccacactgcttccaagt 240
 |||
 Db 5929 AGCTCCGGAGGCTTACTCTAGAGAAAGAGACTTGCAGATTCCACATCGCTTCAAGTGC 5988
 QY 241 gagattcagctatgttagagcgcaacgctgcaggaacactcgagactcggttgaagagac 300
 |||
 Db 5989 GAGATTAGTCTATGTTAGAGCGCACCATCATGAGAAACCTCGAGACTCGGTGGAAGGCGC 6048
 QY 301 ctatctcgaactttgacacacacacggcgctggaagtcgaaggaalcccaacttga 360
 |||
 Db 6049 CTTATCTGCTACTTTTACACACACACACGCGCTGTGAAGTCTGAGAGAAATCTCCACTCGGA 6108
 QY 361 tccatgatcccaagcttaagcttgcgccacactcccgactcggttgagagccgaaaaa 420
 |||

Db 6109 TCCATGCAATCCACGTTAAAGCTGGCGCCACCTCCCGACTCGGGGTGGAGAGCCCAAAA 6168
 QY 421 ctgaagaatcccttaagcttcgctccatccgcttgcttactactaacaataactcc 480
 |||
 Db 6169 CTGAGAAATCCCTTAAGCTTCGCTTCATACGCGCTGTTCTTACTTAACATAAATCTCCC 6228
 QY 481 caggcgaagtaagtaaacgcttataagcgcctcgaaccccccaataagccttaccctacc 540
 |||
 Db 6229 CAGGCCAGTAGTAACGCGCTTATAGACAGCTCGAACCCCAAGACCTTTATCCCTTACG 6288
 QY 541 tggcttatattgacccctatacagggctgcaactgtaaaataagcagcgagtgctctc 600
 |||
 Db 6289 TGGCTATATTATGACCCCTGATACGGGTGTGCACTGTAAATAGACCTCGAGGTGTTCTCCT 6348
 QY 601 aagagacacttgctgagcctgaactgcatcttgctcctcgatttgataaccgcgtgtaa 660
 |||
 Db 6349 ACAGGACACTGGTGGCTGCACTGCAATTTCTGCTCCGATGATTAACCCGCTGTAAA 6408
 QY 661 agcaaacctcccaactagtcgtagttagttagttagttagttagttagttagttagttag 720
 |||
 Db 6409 ACACACACTCCACACCTAGTCCGTAGTTAGGTTCTATGCTGCCACAGGACACAGAA 6468
 QY 721 gagaataactgtgaggttcttgaggaaactctctgtaagagatggaagctgcaactcc 780
 |||
 Db 6469 GAGAAATACTGTGGGGGTCTGGGGAATCCTTCTGAGAGATGAGAGCTGGGTACCTCC 6528
 QY 781 aacagatgaagctggaatgagcgcgactcctcctcgaagcagggttaaatctcctgtc 840
 |||
 Db 6529 AACGATGAGGACTGGGAAAGGCGGATCTCTCCAGGACCGGATTAATTTCTCTTTTTC 6588
 QY 841 aactccggcgagcgagacagacaaaatgaataacataataagataagagcgcgcctcc 900
 |||
 Db 6589 AATTCGCGCGCGGCAAGTACAAAGTATGATGAACATATATAAGATTAAGCTGCTCCCA 6648
 QY 901 tcaagattagattatcctaaagaataagttcaactgaaagaaagaaacagaaataatcaa 960
 |||
 Db 6649 TCAGACTTATGATTACTTAAGATTAAGATTAAGTTCCTACTGAAAAAGAAACAGAAAAATATCAA 6708
 QY 961 aagtgatataatggtatgagcgtggggaatagtttataataatagcgggggagcaggg 1020
 |||
 Db 6709 AAGTGATTAATAGTATGATGAGCTGGGCAATAGTTTATTAATATAGCGGGGACACAGG 6768
 QY 1021 tccacttaacatccgcttagagatagagacgggagaaacccccctgtgcaatggga 1080
 |||
 Db 6769 TCCACTTTTACCATTTGCGCTTAGATAGAGACGGGGACAGAAACCCCTGTGGCACTGGGA 6828
 QY 1081 cccgataaagtaactgctgtaaacaggggcccgccctgtagcacaactctgctg 1140
 |||
 Db 6829 CCCGATTAAGTACTGGCTGAGAACAGGGGCCCCCGGCTGAGGCCACCGCATTAACCTTGGC 6888
 QY 1141 gtgccccaatataactcgtcgtgcgcgtgacataaacaacgcgctagagaacagtaacc 1200
 |||
 Db 6889 GTGCCCAATTAACTCGCTGGCTGACATTAACACAGCGGCTTAGCANAGTACCACT 6948
 QY 1201 ggaattgattctaccacaacgcctagaaactcccaagtgcttctcgtttaagacaggaag 1260
 |||
 Db 6949 GGAATTGATTCCTACCAACAGCGCTAGAAATCCCAAGGTGTTCTGTTAAGACAGAGACAG 7008
 QY 1261 agactcttcagtcacacagggagcttccaagcactcaactccaacgacactgtagcc 1320
 |||
 Db 7009 AGACTCTTCAGTCTATCATCAGGAGACTTTCGCAAGCATCACTCCACCGACCTGATGCC 7068
 QY 1321 actcttccttgctgcttctcctcctcgaagggcctcctttaaagaggaatgagctaa 1380
 |||
 Db 7069 ACTTCTCTTGTGGCTTGTCTATCTCAGGGCTCTCTTATTAAGAGGAGTGGCTTAA 7128
 QY 1381 gaaagaaaatcaatgctacacaaagagacatagaatcaaatgatactgagggctccgaat 1440
 |||
 Db 7129 GAAGGAAATTAATATGTGACCAANAAGCATATGAATATCATGTGAGATGGGGTCCGGAAT 7188
 QY 1441 aagcttaacctcaactgaagtttcggggaaggagacatgcatatagaagaaagctcccaatcc 1500
 |||
 Db 7189 AAGCTTACCTCTACTGAGATTTCGCGGAAAGGAGCATGATAGGAAAGCTCCCATCC 7248

QY 1501 caccacaaccttgcctagtagtgcgtggttaataagcagcctcagaataatcagatttta 1560
 Db 7249 CACCAACACCTTTGGCAATAGTACTGTGTTATGAGCAGGCGCTCAGAAATACATATTTA 7308
 QY 1561 gtacctgttatacaggtggtgacatgcaatagctggttaaccctctgttccacc 1620
 Db 7309 GTACTGGTATTAACAGGTGTGGCATGCAATACGTGGTTAAACCCCTGTGTTCCACC 7368
 QY 1621 tcaagcttcaacccaalcccaagaattgtgtlcaatggtlccaaatcgtlccccagatgac 1680
 Db 7369 TCAGCTTCAACCAATCCAAAGATTTCGTGTCAATGGTCCAAATCGTCCCGCAGGTGAC 7428
 QY 1681 taacactcaggaaggtggtccctgataataataactagctggtataacagcaaaaaa 1740
 Db 7429 TACCACTCTGAGGAAGTGTCTGATGAAATATGACTATCGGTATTAACCGCAAAAAGA 7488
 QY 1741 gaaccgfatcccttaaccctagctgttaattgctcgartagaagcagcgttggtgata 1800
 Db 7489 GAACCGATACCTTACCTAGCTGTATGCTCGGATTAGGAGACGCGCTTGCGCTAGA 7548
 QY 1801 acaaggagacagctgcccctgatacagaagacacagcagctagaagaagactgtgtgagta 1860
 Db 7549 ACAGGAGACAGTCCCTCATATCAGACAGACAGCAGCTAGAGAAAGACTTGTGAGCTA 7608
 QY 1861 catcgagccatgacagaagaatctccagcctlaaagggtctgttaagaacctagaagag 1920
 Db 7609 CATCGGGCCATGACAGAAGATCTCGAGCCTTAGAGGAGTGTGTAGCAACCTTAGAAGAG 7668
 QY 1921 tcccgaactctctgtcgaaggtgtctacagaacagcagaggtgatacctcgtt 1980
 Db 7669 TCCCTGACTCTTCTGTGAATGTTGTACAGAACCGAGGCGATTAGATCTGCTTT 7728
 QY 1981 ctaagaagaagtggtgtatgtgcagcctlaaagaagaatgtgtcttctatgtagatc 2040
 Db 7729 CTAAAGAGAAGTGGGTGTGTGTCACCTTAAAGAGAATGTGCTTATGTATGATAC 7788
 QY 2041 taagaagccataagaagactccatgaagaagcttagaanaaagcttagaagcgttcgaag 2100
 Db 7789 TTAGGAGCCATAGAGACTCCCTAGCAAGCTTACAGAAAGGTTAGAGAGCGTGAAGG 7848
 QY 2101 gaaagaagaagctgacagaaggtgtgttgaagatgtgtcaacaggtctcctgtgatacc 2160
 Db 7849 GAAAGAGAGGTGACAGAGGTTGTTGAAGATGTGTCAACAGGTCTCCTTGATGAGC 7908
 QY 2161 accgtgtctctgtctcgaaggggcccctagtagtgcgtctcgttcttactcaagtgg 2220
 Db 7909 ACCCTGCTTCTGCTGTGACGGGGGCCCTAGTAGTCTGCTTACTTACAGTTGGG 7968
 QY 2221 ccttgttaataataggtgtgttccttgttagaagaagtagtagtcaatccagatc 2280
 Db 7969 CCTTGCTTAATTAATAGGTTTGTGCTTGTGTAGAGAACGATGATGCTCAGATC 8028
 QY 2281 atgttactaagcaacagtagcaagcctctgagccaaagagaactgacctagacct 2340
 Db 8029 ATGCTACTTAGGCAACAGTACCAAGGCTTCTGAGCCAGAGAAATCCTTACCTTACGCT 8088
 QY 2341 tcccagttcttaagattagaactataaacaagaagaagttgaggaatgaagaatgaaga 2400
 Db 8089 TCCCACTTCTAAGATTAGACTATTTAACAAGACAAAGATGGGGAATGGAAGATGAANA 8148
 QY 2401 tgaacttaaacctcccaaaacccaagaagttlaataaaagctlaaatgccccgaatt 2460
 Db 8149 TGAACCTTAACCTCCCAAGAACCCAGAGATTAAATGAAGCTTAATGAATGCCCGCAATT 8208
 QY 2461 acagaccgtggtgtgcagtaataagtagaaggttcaacttctatgttccagggc 2520
 Db 8209 CCAAGACCTGTGCTGTGCTAGTAATAGTAGAAGGTCACTGTGTATTTGTTCCAAAGC 8268
 QY 2521 ctgtctctcgtgacatagtagaataacagaagaatggttgaactacgttcttctgatt 2580
 Db 8269 CTGCTATCTGTGCTTAAGTAGATTAACAGGAATGAGTTGACTTAATCGCTTATCTGATT 8328

QY 2581 ctgtaaacactgactggcacatagaagaattgattlacatagtagacacctagtagacta 2640
 Db 8329 CTGTAAACACTGACTGGCACCATAGAAATGTGATTACACTTGTACAGCCCTAGTACTA 8388
 QY 2641 tctcaactgcaatctgtcaactcgtccaggaagcccaagatagtgagacctccagagcta 2700
 Db 8389 TCTGAACCTGCAATCTGTGACTGTGCGCAGTAGGCCACGAGAGAGCGACCTCCGAGCTA 8448
 QY 2701 tttaaaatgattgtgtccagcagcagcagcgtctgattttaaattgattgctcag 2760
 Db 8449 TTTTAAATGATGTGTCCACGAGCGCGGCTCTGCATATTTAAATGATTGTGTCATG 8508
 QY 2761 gagcgcggtctcgcattatllaaatgattgtgtgacagcagcgttgtgtgaa 2820
 Db 8509 GAGCGCGGCTCTCGATTTTAAATGATTGTGTATGAGACGCACAGCGCTTGTGTGAA 8568
 QY 2821 ccccaataaaagcgtgccgaattccgaactcgtgggagcgcagcttacccttgctgtg 2880
 Db 8569 CCCCATTAAGCTGTCCCGCATTCGCACTCGGCGCGCAGCTCTTAACCTTCGTGCTG 8628
 QY 2881 taagactgtggcccccagcagcgttgaataaaatcctctgtgttgcatacaaa 2938
 Db 8629 TACGACTGTGGGCCACAGCGCGCTTGATTAATAAATCTTCTGTTTCATCAAGA 8686

RESULT 5

LOCUS AX002804 3482 bp DNA PAT 21-AUG-2000
 DEFINITION Sequence 3 from Patent WO9853104.
 ACCESSION AX002804
 VERSION AX002804.1 GI:9885132

KEYWORDS

SOURCE porcine endogenous retrovirus.
 ORGANISM porcine endogenous retrovirus.

REFERENCE 1 (bases 1 to 3482)
 Stoye, J.P. and Weiss, R.A.
 Detection of retroviral subtypes based upon envelope specific sequences
 JOURNAL MEDICAL RES COUNCIL (GB): STOYE JONATHAN PAUL (GB)

FEATURES

source
 1. 3482
 /organism="porcine endogenous retrovirus"
 /db_xref="taxon:61673"
 BASE COUNT 927 a 854 c 867 g 834 t

ORIGIN

Query Match 97.7%; Score 2889.4; DB 6; Length 3482;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 2913; Conservative 0; Mismatches 21; Indels 1; Gaps 1;

QY 1 tgcctttaggttagaagaacccctggaagcttgggtgacccctatgaattgtct 60
 Db 548 TGCCTTTAGGTTAGGAACAACCCCTGACAGTTTGGCTGACCCCTATGAATTTGCTCT 607
 QY 61 acggggagacccccccgctgtgtagaatgtctctgtacatagtgctgattgtctt 120
 Db 608 ACGGGGAGACCCCCCGCTGTGTAGCAAAATGCTTCTGTACATATGATGATGCTGCTT 667
 QY 121 cccagccttgtctcctagctcaagcgcgtcgaagtgagtggaagcaagcgtgtgaagc 180
 Db 668 CCCAGCTCTGTCTCTAGGCTCAAGGCGCTCGAGTGGGTGAGCAACGACGCTGGAAGC 727
 QY 181 agcttcggagagccttaccagaagaagagagcctgcaagttccacatgcttccaagt 240
 Db 728 AGCTCCGGAGGCTTACTCAGAGAGAGACTTGCAGATTCCACATGCTTCCAAAGTG 787
 QY 241 gagatcagctatgttagacgccacgtgcaagaaactcgaagactcgtgtgaagggac 300
 Db 788 GAGATTTCAGTCTATGTATGAGCCACCGTGCAGGAACCTCGAGACTCGTGGTGAAGGCC 847

QY	301	cttatcttcgtaacttttgaccacacacagcgtgtgaagaatcgaatctccacttga	360
Db	848	CTTATCTCGTAACCTTTGACCAACACCAAGCGCTGTGAAGTGCAGGAATCTCCACCTGGA	907
QY	361	tccatgatccacaggttaaacgttggccgaactcccgactcggggttgaagagccgaaaga	420
Db	908	TCCATGATCCCAAGTTAAAGTGGCGCACCTCCGACTGGGGTGAAGAGCCGAAAGA	967
QY	421	ctgagaatcccccctaagcttcgcctccatcgctcgtgtcttactactcaaatlaactccc	480
Db	968	CTGAGAAATCCCTTAAACCTTCGCTCCATCGCGCTGCTTCTTACTCTAACAAATTAACCTCC	1027
QY	481	cagggccagtagtaaacgcccctatagaagaatcgaacccccatagaaccttataccctaac	540
Db	1028	CAGGCCAGTAGTAAGCCCTTATAGACAGCTGCAACCCCATAGACCTTATATCCCTTACC	1087
QY	541	tggctgattatitgacctgtatagcgytgcactgttaaatgacctcgaagtgtgttcctc	600
Db	1088	TGCTGTATTTATGACCCCTGATAGGGGTGTACTGTAAATAGCACTGAGAGTGTTCCTCT	1147
QY	601	agaggaacctgtgtgtgcctgaacatgcattcttcgcctccogattgataaaccgcgtgttaa	660
Db	1148	AGAGGACCACTGGGGCTGTAACTGCATTTCTGCCTCGATGTATTAACCCCGCTGTAA	1207
QY	661	agacacccctcccaactagtcgtagttagttatggcttattgtctgcccagcagaagaaa	720
Db	1208	AGCACACCTCCCAACTAGTCCCTAGTTATGGGTCTATTGTCTGCCAGCACAGAGAA	1267
QY	721	gagaaatactgttgggggttcctggggaatcctctctgtagagatgtgcgtcacctcc	780
Db	1268	GAGAAATACMTGGGGGCTTCTGGGGAATCCTTGTAAGAGATGAGAGCTGCCTCACTCC	1327
QY	781	aacgaatggagactcggaaatggccgatactctctccagagccggtaaatcttccttgtc	840
Db	1328	AAACGATGGAGACTGGAAATGGCCGATCTCTCCAGACCGGGTAAATTTCTCTTTGTC	1387
QY	841	aattccggccggcggcagatcacaaatgatgaactatacaagaatgaagcgtcctccca	900
Db	1388	AATTCGGCGCCGGGCAAGTACAAAGATGAANACTATATAAGATTAAGCTGTCTCCCA	1447
QY	901	tcaagcttagattacttaagaataagtttctactgaaaaaagaaacaggaataattcaa	960
Db	1448	TCAGACTTAAGATTACTTAAGATAAGTTTCACTGAAAAAGGAAACAGAAATAATTC	1507
QY	961	aagtgataaatgtatagctcgtgggaatagtttttaataatatgtggcgggaacagag	1020
Db	1508	AAGTGATTAAGATGATAGCTGGGGATAGTTTTTATTAATATAGCGGGGACAGAGG	1567
QY	1021	tccactttaacatctgcctcttaggaatagacggygagacaaaccccctgtgcaatlyga	1080
Db	1568	TCCACTTTAAACCATTTGCGCTTAAGATAGAGACGGGACAGAACCCCTGTGCAATGGGA	1627
QY	1081	cccgataaagtaactgtgcgtgaacaaagggcccccggccctggagccacgcgataacttgcg	1140
Db	1628	CCCGATAAAGTACTGTGCTGAACAGGGGCCCGCGCTGGAGACCCCATTAATCTTGCG	1687
QY	1141	gtgcccaatlaaacctcgtcgtcggtctgacataacacagccgctcaggacaagtaacca	1200
Db	1688	GTGCCCAATTAACCTCGCTGGGCTGTACATTAACACAGCCGGCTTAGACAGAGTACCACT	1747
QY	1201	ggattgattctctacaaacagccttagaaaactcccaggtgttctcgtttaaagacagagag	1260
Db	1748	GGATTGATTCTCTACCAACACCCCTAGAAACTCCAGGTGTTCCTGTTTAACACAGACAG	1807
QY	1261	agactcttcaatctcatccagggagccttccaagccatcaactccaacgaccctgatgccc	1320
Db	1808	AGACTCTCACTCATCTCAGAGAGCTTTCCAAAGCCATCAACTCCACCGACCTGATGCCC	1867
QY	1321	actctctctgtgtgcttgtctatcctcagagccctccattatagaagggatgtgctaaa	1380
Db	1868	ACTCTCTCTGTGTGCTTGTCTATCCCAAGGCGCTCCTTATATAGAGGATGSGCTAAA	1927
QY	1381	gaaagaaatctcaatgtgacccaagagcatagaatatgaatgaatgaggggtccggaaat	1440
Db	1928	GAAAGAAATTTCAATTGTGACCCAAGAGCATATGAAATTAATGTATATGGGGTCCCGAAAT	1987
QY	1441	aagcttaacctactgaagttccgggaagggagatgcaatagaagaaagctcccacatcc	1500
Db	1988	AAGCTTAACCTCACTGAAGTTTCCGGGAAGGGGACATGATAGGAAAGTCCGCCCATTC	2047
QY	1501	caccaaaccttggctatagatctgtgttattagaagagccctcagaataatcagattata	1560
Db	2048	CACCAACACCTTTGCTATATGATAGCTGTGTTATGAGAGGCGTCACGAAATACGATATTA	2107
QY	1561	gtacccgttataaacaaggtgtgtggcagatgcaatactgtgtlaacccccctgttccacc	1620
Db	2108	GTAACGTGTATTAAGAGGTGTGGGATGCAATACGTGGTTAACCCCTGTGTTCCACC	2167
QY	1621	tcaagtcttaaccaatcccaagaattgtgtgtcatatgtgtccaaatcgtlccccagatgtac	1680
Db	2168	TCAGTCTTCAACCAATCCAAAGATTTCTGTGATAGGTGTCGCAATATGCTCCCGCAGGTATC	2227
QY	1681	taccatccctgaagaaagtgtgcctctgtatgaatatgacatcogtataaccgaccaaaga	1740
Db	2228	TACCATCTGAGGAAGTGTGCTTGTATGATATGACTATGGGTATTAACGACCAAAAAGA	2287
QY	1741	gaacccgtatcccttaacctagctgttaatgtcgtgatalaggaagcgcgtltggcgtlaga	1800
Db	2288	GAACCCGTATCCCTTACCTTACCTAGCTTAATGCTGATGATTAAGGACGCGCTGTGCGTAA	2347
QY	1801	acagggagacagctgtgcctgtgtacaaagacacacagcagctagagaagaagacttgttagcta	1860
Db	2348	ACAGGGACAGCTGCCTCTGATCAAGAGACACACAGCACTAGAGAAAGACTTGGTGAAGTA	2407
QY	1861	catgcygccaatgacagaagaatccagagccttaaaagagctgtgttagcaacctagaagag	1920
Db	2408	CATGCGCCCATGACAGAAGATCTCGAGCCTTAGAGGATCTGTTAGCAACTAGAAAGC	2467
QY	1921	tccctgactcttctgtctgaagtgtgttctacgaacccggagggatltagatctgctgtt	1980
Db	2468	TCCCTGACTCTTGTGTGAAGTGTGTCATGATGTCACCAACCGAGGGGATTAATGATCTGCTT	2527
QY	1981	ctaaagaaaggtggtgtatgtgtgcaagccttaaaagaagaatgtgtgcttctatgttaaac	2040
Db	2528	CTAAGAGAAAGTGGGTATGTGTCACACTTAAAGAAAGATGTGCTTCTATGTAAATCAC	2587
QY	2041	tcaagagccatcagaagctcoaagaagaagcttagaaaaaaglttagagagcgtlccgaag	2100
Db	2588	TCAGAGGACATCAGAGACTCCATGTAGCAAGCTTAAGAAAGGTAGAGAGCGCTGCAAGG	2647
QY	2101	gaaagagagcctgacaggggtgtgtttaaagaatgtgttcaacaagtctcccttggatgcc	2160
Db	2648	GAAGAAGAGGCTGACCAAGGGTGTGTTAAGAGATGTTCACAGAGTCTCTTGGATGAC	2707
QY	2161	aacctgtcttctgcctcgaagggcccttagttagtccctgcctcttacttacaatttgg	2220
Db	2708	ACCCTGCTTTCGTGCTGACGGGACCCCTAGATGCTGTCTCTGTACTTAACATTTGG	2767
QY	2221	ccttgcttaattaatagtttgttgccttgtttagaagaacagatgtgcagtlccagatc	2280
Db	2768	CT	

SOURCE porcine endogenous retrovirus.
ORGANISM porcine endogenous retrovirus
REFERENCE Viruses; Retroviral viruses; Mammalian type C
AUTHORS 1 (bases 1 to 7808)
TITLE 1 (bases 1 to 7808)
JOURNAL Cauberna, F., Fischer, N., Bolter, K., Kurth, R. and Tonjes, R.R.
MEDLINE Establishment and characterization of molecular clones of porcine
REFERENCE endogenous retroviruses replicating on human cells
J. Virol. 74 (9), 4028-4038 (2000)
20219394
2 (bases 1 to 7808)
AUTHORS Tonjes, R.R.
TITLE Direct Submission
JOURNAL Submitted (30-MAR-1998) R.R. Tonjes, Paul-Ehrlich Institut,
Paul-Ehrlich Str. 51-59, P.O. Box 1740, D-63207 Langen, FRG
FEATURES
1. 7808
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
/tissue_type="kidney"
/cell_line="PK15"
260. 1834
/gene="gag"
260. 1834
/gene="gag"
/codon_start=1
/protein_id="CAA76581.1"
/db_xref="GI:3005986"
/translat="MGCTTTPSLTLHWTVRSRAHNLVSQKKGMQTCASEMP
TFDVPWSESTENETILAAKAIIFQGTGSSHPDEPYLLTMQDAEDPPVACFWLN
PKRPGRPIALAEKNSHAEKVEPSPIRYPEIEPTWPEPPVPPVACQWAGV
PSAPCAPVPEPAAGTSSRGATERTDEIILATRTGPPGCGOLLOMYFESS
ADLYNKTNNPESDEPRLTGIVSISMHOPTMDCOLLOLTFTPEERRILLES
RKVPGADRPQLOLNETMDGPIRLGMDYTAGRESLKITRYAIVAGLGAARRP
TNLAKEYVQGNBPSPVFLERLMAFRRTPEPDSIAQASVALAIGQALDIR
KTLQRLGLOEALDLVREAEKAYGRTTEKEQRKEEREERDRQENLT
KTLAAVEGKSSEREERDRKTRSGPDSGNLTPLDKDCAYCKEKGHWARNCPK
KGNKGPVLALEDKD"
2120. 4726
/gene="pol"
2120. 4726
/gene="pol"
/codon_start=1
/protein_id="CAA76582.1"
/db_xref="GI:3005987"
/translat="WGAQISFEQGRREVSVNKPITVLTLDIDEXRLYSPQKPPDQ
IOSWLEQFOQAETAAGMGLAQVPOVQLKASATPVSRQYPLSRAREIMPHVQ
RLIOGILVPOSPWNTPLLPVAKGTNDYRPVODLEVNKRVODIHTVPPRYLS
ALPBERNWTVDLKDAPFCLRLHPTSOPLFAFERMDGTGTGOLTMRLPOGKNS
PTIPEDALHDLNFRIOHPOVTLIOYDNLIDAGATODCEGKALLILESDGYR
ASAKKAOICREVTYLGYSLRGQORMLTBAKKTIVQIDAPTAAOVEEFGTAFGR
LMTPGATLAAPLYPLTKEGFSWAPEOKAFDAIKKALLSAPALADVTKPTLY
YDERGVANGVLTQTLGPKRPVAVLSKLDVAGWGVCEKAIAAVILKADAKLT
LQONTIVIALENTIVRQPDMMNTAMHYOSLLTERVTFAPPALNDATLPE
ETDEPETHDCHOLLIEGVGRKLDIDPLTGEVLTWPTDSSSYVEGGRMAGAYOG
THTIASSLPEGSAOKAELMALTOALRAEKSINITDSSYARATRHVAIYKOR
GLTISAGREIKKKEELISLELHPKRALIHCGGHOKAKDLISGQOMADRAKQA
AOAVNLPITTEPRAPERROYTLEDWCEIKIDQSETPGTCYSTGKELPKES
LEBYOOLHRLTGLTKHLOLVRTSPVHLRLPGVADSVKHCVCOLVNNPSITP
GKRLGSHPGAHMEVDTEVKKPAKGNKYLIVFVDFSGWEAYPTKKETSTVAKY
WRKFFRDELYLH"
5294. 7267
/gene="env"
5294. 7267
/gene="env"
/codon_start=1
/protein_id="CAA76583.1"
/db_xref="GI:3005988"
/translat="MPTLSMRHLPTRGCEPKRLRILPSPASIAFWLTTLTPQASAK
RLIDSSNPHRPLSLTWLIIDPDTGVNTSGVAPRGTMWDELHPCRLINPAVSTP

PNLVRSYGYCCPGTEKEKRYCGGSEESFCRRMCSVTNSDGMKMPISLDRKFSFVN
SGPKRYKMYLYKDKSCSPSDLYLKIIFTEKGKENTONKINGSMWIVRYKYGGA
GSTLIRLRIENGTEPPVAGVDKVLADGPPALPEPOTLRPIOTPSN
GTTGILPTNTPNSPVKQGRALFSLICCAFQAINSTDPAJSCMLCISGSPRY
BMAKEGFNWTKEKRNCTWKSRRKLITLVESSGCTGTGAPSHOHLCSTVVEO
ASENOYLVPGRNRMWACNTGLTPCVSSVFNOQKFCVWQIVPVYVHREPVLEDE
DYRNRPKREPVSLTAVMLIGTAVGCTALITGPQOLEGIGLHAMVEDLR
ALIEESVNLIESLTSLSSEVVLONRGIDLTLREGICLALKRECCFVYDHSGLRDS
MSKRELERREERERADQWMEEGFNFSPPMPTTILSLATGPVLVLLLVGCLIN
RVAFVRREAVQOIMVLRQOYGLISGGERLDL"
BASE COUNT 2077 a 1971 c 2047 g 1713 t
ORIGIN
Query Match 96.1%; Score 2839.6; DB 14; Length 7808;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2854; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 1 tgccttttaggttagaaaccccttgacagcttggagctgaaccccttaagatgctct 60
DB 4931 TGCCTTTTAGGGTTAGAACACCCCTGACAGCTTTGGGCTGACCCCTATGAAATGCTCT 4990
QY 61 aaggggagaccccccccgcttggtagaaattgctctgtacataagtgctgctgtctt 120
DB 4991 ACGGGGAGACCCCCCGCTTGTAGAAATGCTTGTACATAGTGTGCTGCTGCTTT 5050
QY 121 ccagagccttgcctctaagctcaagggcctgaagtgagtgaggaagagctggaagc 180
DB 5051 CCCAGGCTGTGCTCTAGGCTCAAGGCGCTGAGTGGGTGAGCAGAGCGTGGAAAGC 5110
QY 181 agctccggagagcccttaagtagagagagagagcttcaagttccacatcgctccaaagt 240
DB 5111 AGCTCCGGAGGCGCTACTAGAGAGAGAGAGACTTCCAGTTCCATGCTTCCAAAGTGG 5170
QY 241 gggatcagctcatgttagaagccacgctgcaagaaacctcgagctcgttgaaaggac 300
DB 5171 GAGATTCACTATGTTTATGACCCGACCGCTGAGAAACCTCGAGACTGCTGTAAGAGGCGC 5230
QY 301 cctatcctgactctttgaccacacaaaggcctgtaaaagtcgaagaatctccactgga 360
DB 5231 CTTATCTGCTACTTTTGACACACACAGGCGCTGTAAGTCCAAAGAAATCTCCACTGGA 5290
QY 361 tccatgcattccacgtttaagctgagccacccctccagactcgggagagagcgaagaa 420
DB 5291 TCCATGCATCCACAGCTTAAGCTGGGGGCCACCTCCCGACTCGGGGTGGAGAGCGGAAAGA 5350
QY 421 ctgagaatcccttaagcttgctccatcgctcgttgcttccactactaacaataactccc 480
DB 5351 CTGAGAATCCCTTAAGCTTGCGCCATCGCCTGTTCTTACTCTTAACATAACTCTCC 5410
QY 481 cagggccagtagaaacgctttaagaagcttcgaacccccataagaccttaccctaac 540
DB 5411 CAGGCGCAGTAGTAACGCTTATVAGACAGCTGGAACCCCAATAGACCTTTATCCCTTACG 5470
QY 541 tggcgtatattgaccccgatagcgggtgtaactgtaaaatagaacccgaggtgtgctct 600
DB 5471 TGGCGTATTATTGACCCCGATAGCGGGTGTCACTGTAATAGCACTCGAGGTTGCTCTCT 5530
QY 601 agaggaacacctggtgctgcaactgcatcttgctccgatgatgaatacccgctgttaa 660
DB 5531 AGAGGACACCTGGTGGCTGGAAGCTGATTTGCGCTCCGATGATTAACCCGCTGTAAA 5590
QY 661 agcacacctcccaactgctcgttagtgatagtggtctctatgctcccaaggacagagaaa 720
DB 5591 AGCACACCTCCCAACCTVAGTCCGTAATTAAGGTTCTATTTCTGTCCGACGACAGAGAAA 5650
QY 721 gagaatactgtgaggtctcgggaaacccctctgtaagagatggaagctgcgtacacctc 780
DB 5651 GAGAAATACTGTGGGGGTTCTGAGGAATCTTCTGTAGGAATGAGAGCTGCTACACTCC 5710
QY 781 aacgaatgagacgtgaaatgagcagatctctctcgaagacgggttaaatctctctgtgc 840
DB 5711 AACGATGAGAGACTGGAATGAGCGCATCTCTCCAGAGACCGGTAATAATCTCTCTTGTGC 5770

SOURCE protein.
ORGANISM porcine endogenous retrovirus.
REFERENCE Viruses; Retroviridae: Mammalian type C
AUTHORS (bases 1 to 8918)
TITLE Caudeana, F., Fischer, N., Boller, K., Kurth, R. and Tonjes, R.R.
Establishment and characterization of molecular clones of porcine
endogenous retroviruses replicating on human cells
JOURNAL J. Virol. 74 (9), 4028-4038 (2000)
MEDLINE 20219394
REFERENCE 2 (bases 1 to 8918)
AUTHORS Tonjes, R.R.
TITLE Direct Submission
JOURNAL Submitted (04-MAY-1999) Tonjes R.R., Medical Biotechnology,
Paul Ehrlich-Institut, Paul Ehrlich-Strasse 51-59, Hessen, D-63225
Langen, GERMANY

FEATURES
source location/Ouallifiers
1..8918
/organism="porcine endogenous retrovirus"
/proviral
/strain="type C"
/db_xref="taxon:61673"
/clone="33"
/note="grown in PK15 (ATCC No. CCL-33) cell line"
1..707
545
/note="cap site"
710..727
/note="g1y4"
1154..2725
/gene="gag"
1154..2725
/gene="gag"
/function="retroviral capsid"
/product="gag protein"
join(1154..2725,2729..6313,6189..8162)
/note="viral polypeptide"
/codon_start=1
/protein_id="CAB5339.1"
/db_xref="GI:6688946"
/translat="MGCTVTPPLSLTLDHMTVEVSRANLTVQKKGPMOTFCASEMP
TFDVGMPSEGFENSEILIAKAIIFQGPSHPPOEYILTMWDLADPPKPMWLN
KPRKPGRIILGCKNKHSAKVEPSRPIPELEPTPEPOVPPVPAQAVAG
PSAPPGAVVEGPAAGTRSRGATPEPTDELAIIPLTYTPMAGGLOLOLQWPPSS
ADLNKMTNHPSEEDPQRLGLVESLMSHQPWDCDOLLOTLFTTEREKLLEBA
RNKAVGDEGPTOLQNEIDMGFLTRGMDYNTAEGRESLKYQALVAGLRASRP
TNLAKVEVMGPPPEPVELERLEMAFRFTPDPTSEKQKASVALATIGOSLDIR
KTLQRLGLOAELRLDYREAEKYYRRETEKEKREERERERDRROENLIT
KTLAAVEGKSRERERDRFRIKSGPROSNGTIGRTLDNOCACKEKGMHNCPEK
KGNKPKYITILEEKDGRGSDPLPEPRVITLKVGQVVELEVDGASHVLLDPLKTL
KEKSWMGATGQROYFWITRTYDLVAGRTSHFLVIPCVPFLRLDLTKMAOI
SPEGRPEVSNKPIYVLTLDDEYRLSPQVPODIOQWLEOPQMAETAGMG
LAKOVPOVIOLKASATPEVSVQYRLSREABEGIMPVORLIQGLVPOSPWNTPL
LPVRPGNDYRPODLREVNKRIVODIHPVNPYNLISLIPERMYVTLIDKDAF
CLRHPISOPLEAFEMRDPGRTGOLMTPLPGCFKNSPTIDEALHRLDRANRICH
POYTLLOYDOLLAGAKKODICEGTAKLILELSDIGRASAKAOICREAVTIGYS
LNGGRWLTERRKRTVOIIPATTAKOVREFLTAGTCRIMTPEFATLAPLYLITE
KGESWAEHOKADAIKALLSAPALAPLVNTPETLYDERGVARGVLTQVLGFW
RRPVAVLSKLDPAVAGWPCLKAIAVAVALIKADKLTIGONTTVAIPAHLENTVO
PPDRMTNARMTHYOSLTLERTFAPPAALNPATLLPEETDEFTDCHOLETENG
VRKLDLIDPLTGEVLTWFTDGSYSVEKRMAGAAVVDGHTIWASSLESTQKAE
LMALTOALRIAEKSIINYDSPAFAAHYAHYORGLITSAGREIKKREILSL
LEALHPKRLAIIHCRGOKAKDLISRONADVAQOANAVMLPIETPKAPPR
ROYLEDMEETIKIDQSETEPGTQYISYGEELPHKBGLTVOQIHLTHLTKHIO
OLVRTSPYHLRLPGVADSVKHCVPOLVANASRLPQGRKLGSHPGAWEDVTE
VPAKYGKYLIVVDPSGWEAVPTKKESTAVAKKILEIIPRIGIPVISDNG
PAFVAVOSGLAKIIGIDMKLHCARPOSSQVGRMNTIETLTKTTEGINDWA
ILPVLFEVRNTPGOGLPYELLIGGPPYELVIAVSASAVILSOPLESKLAEW
ROARKOLREAVSGRGLQVPHRVOGYSYVRHRRGNLFTBRKGPVLTITPTAV
KVEGISTVHTSHVAKLAPPSDGSMAETEDPLKRLRLRVIPYNNNSPGCIHTPTLS
RHLPTRGEPKRLRIPISFASIAWFLTLTTPQASSKRLIDSNPHRLSLTWLIDP

DTGVTNSTRGVAPRGCTWPELHFLRLINPAVKSTPPNLYRSYGYCCPTEKEXYC
GDGSEFCRMSCVTSNDGDKWPISTLDRKFSFVNSGPKYKWKLYKDKSCSPSD
LDYKISTFEKQENIKWIMGMSGIVFYKGGAGSTLTIRLTETGPEVPAVG
DKVLAEOGPALPEPNNLPVOLTSLRPIIDIOSPSNGTGLIPNPRNSGPVKT
GORLSILIOGAFQALNSNDPATSCTWCLTSGPPYEGMAKEGFVNTKEHNOCTW
GSRNKLITLVSNGKTCIGKAPPSHOHCISYVYEOASENOYLPVGNRRMACNGL
TPCVSTSVFNOSKDFCVMOIVPVRYTHPEEVYDELYRINRRRREVSITLAVMS
LGTAVGCTGATLITGPOLKEGSGELHAMTEDRLALESVSNLSEVVL
OORGLDILFTEGGLCAALKECCFYVDHSGAIDMSKILRELRERREREDOGM
FEGMNRSPWMTTLISALGPLVILLITLTPCLINRFVAFRERSAVOIMVLRQO
XGCLISQGETDL"
2729..6313
/gene="pol"
/function="retroviral protease and polymerase"
/product="pol protein"
2729..6313
/gene="pol"
6189..8162
/gene="env"
/function="retroviral envelope"
/product="Env protein"
6189..8162
/gene="env"
LTR
polyA_signal
8811..8816
BASE COUNT 2327 a 2245 c 2334 g 2012 t
ORIGIN

Query Match 95.1%; Score 2811.6; DB 14; Length 8918;
Best Local Similarity 96.6%; Pred. No. 0;
Matches 2914; Conservative 0; Mismatches 24; Indels 78; Gaps 1;

QY 1 tgccttttaggttagaacacccctggaacagtttggtcgtacccctatgaattgctc 60
DB 5826 TGCCTTTTAGGCTTAGAACACCCCTGGACAGTTTGGCGTGACCCCTATGAATTCCTCT 5885
QY 61 acggggagaccccccgcgttgtagaaattgcttcgtacatagtgctgtagtgcctt 120
DB 5886 ACGGGGAGACCCCCCGCTTGTAAGAAATTCCTCTGTACATAGTGCATGATGCTGCTCTT 5945
QY 121 cccagccttctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 180
DB 5946 CCCAGCCTCTCTCTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 6005
QY 181 agctccggagggcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 240
DB 6006 AGCTCCGGAGGCTTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 6065
QY 241 gagatcagctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 300
DB 6066 GAGATCAGCTCTCTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 6125
QY 301 ctatcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 360
DB 6126 CTTATCT 6185
QY 361 tccatgcatcccaagctcctcctcctcctcctcctcctcctcctcctcctcctcctc 420
DB 6186 TCCTATCATCTCCCAAGTTAAGTGGCGCCACTCCCACTGGGGTGAGAGCGGAAGA 6245
QY 421 ctggaatccctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 480
DB 6246 CTGGAATCCCTCTTAACTTGGCTCCATCGCTCGTGTCTCTTACTTCAAAATTAACCC 6305
QY 481 caggccagtagtaagcgcctatagagagcctcgaaccccatagaccttatccctc 540
DB 6306 CAGGCGAGTAGTAACGCTTATAGACAGCTCGAAGCCCATAGACCTTATCCCTTACC 6365
QY 541 tgcctgattatagacccctcctcctcctcctcctcctcctcctcctcctcctcct 600
DB 6366 TGCCTGATTATTAGACCCCTCGATAGCGGGTGACACTTAATAGCACTCGAGGTGCTGCT 6425

QY 601 aagagaccctggtgctgtaacgtcaatctctgctccgatgtatgaacccccgtgttaa 660
| | | | |
Db 6426 AGAGCACCTGCTGGCCGAACTGATTTGCTCCGATGATTAACCCCGTGTAAA 6485
| | | | |
QY 661 agcaacacctccaaactgctcgtatgtatgtgtctatgtctccaaagcaagaata 720
| | | | |
Db 6486 AGCACACTCTCCCAACCTTAGCTCGTAGTATGGGTCTATGCTCTCCACAGGACAGAA 6545
| | | | |
QY 721 gagaatactgtggtgtctggtgaacacctctgtatgagatgtgagctgtcaactcc 780
| | | | |
Db 6546 GAGAAATACTGTGGGATTTCTGGGAAATCTTCTGTAGAGATGTGAGCTGGCTCACTCC 6605
| | | | |
QY 781 aacgatgagacttgaatatgtgcacatctctccaggaacgggttaaatctcctgttc 840
| | | | |
Db 6606 AACCATGAGACTGGAAATGGCCGATCTCTCTCAGAGACCGGATTAATCTCTTGTGTC 6665
| | | | |
QY 841 aattccggcccggaacaataacataatgaataatataaagtaagctgtctccca 900
| | | | |
Db 6666 AATTCCGGCCCGGCAAGTACAAAGTATGAACTATTAAGATTAAGAGCTGCTCCCA 6725
| | | | |
QY 901 tcaagactaga tta tctaaaga taagt ttcactga aaaaagaaacagaaatattcaa 960
| | | | |
Db 6726 TCAGACTTAGATTTCTAAAGATAAGTTTCACTGA AAAAGGAAACAGGAAATATTCAA 6785
| | | | |
QY 961 aagtgtataa tgtatg agctgtgg gaaatgtttttaaataatgtgcgggtgagcaagg 1020
| | | | |
Db 6786 AAGTGTATAATGCTATGAGCTGGGAAATGTTTATTAATATGCGGGGAGCAGCG 6845
| | | | |
QY 1021 tccacttaacatctgctccttaga tagagacgggagacaaacccctgtgcaatgtgga 1080
| | | | |
Db 6846 TCCACTTAACCATTTCCGCTTAGATAGATAGAGAGGGAGCAACCCCTGTGGGAGTGGA 6905
| | | | |
QY 1081 ccggaataaagtaactgtgtaacaggggccccggccccctgtgagccacgcaatactgccc 1140
| | | | |
Db 6906 CCCGATAAAGTACTGCTGAACAGGGGCCCCGCGCTGAGCAGCAGCATTAACCTTGCGC 6965
| | | | |
QY 1141 gtgcccgaatlaactgtcgtcgtgcgcctgacatacaacgcgcgctagcaacagtaacct 1200
| | | | |
Db 6966 GTGCCCCAATTAACCTGCTGGCGGCTGACATTAACACAGTCGCTGACCAACGGGTACCACT 7025
| | | | |
QY 1201 ggalgtatctctaacacaacagctagaacactcccaaggtgttccttlaagagaagagag 1260
| | | | |
Db 7026 GGATGTATCTCTACCAACAGCGCTAGAAACTCCCAAGGTGTCTCTGTAAAGACAGGACAG 7085
| | | | |
QY 1261 agactctcagctcactcagggaggttccaaagcaatcaactccacgacccctgtatgccc 1320
| | | | |
Db 7086 AGACTCTTCAGTCTCATCCAGGAGCTTCCAAAGCCATCAACTCCACCGACCTGTAGGCC 7145
| | | | |
QY 1321 acttctctgtgtgctgtgtctatctcactcagggccctccttataatgagggatgtctaa 1380
| | | | |
Db 7146 ACTTCTTGTGTGGCTTGTGTATCTCTCAGGGCTCTCTATATGAGGGGATGGCTTAAA 7205
| | | | |
QY 1381 gaaagaaatltcaatgtgaccaaagagcatagaatcaatgtatcatgtgggtgtccgaat 1440
| | | | |
Db 7206 GAAGGAAATTCATATGTGACCAAGAGCATAGAAATCATGTATCATGAGGGGTCCGAAAT 7265
| | | | |
QY 1441 aagcttaacctcaactgaagtttccggaaagggagacatgataagaaagctccccatcc 1500
| | | | |
Db 7266 AAGCTTAACCTCTACGTAAGTTTCCGGAAAGGGAGCATGATGAGAAACCTCCCCATCC 7325
| | | | |
QY 1501 caccacaacctgtgctatagtactgtgtttatgagcagggctcagaanaactcaatatta 1560
| | | | |
Db 7326 CACCAACACCTTGTGTATGTACTGTGTATGAGCAGGCGCTCAAAATATCAGTATTTA 7385
| | | | |
QY 1561 gtaccgtgtataacaaggtgtgtgcatgcaactggtgttaacccccgtgttccacc 1620
| | | | |
Db 7386 GTACCGTGTATTAACAGGT 7445
| | | | |
QY 1621 tcaagctcacaacatccaaagatltgtgtgtcaatgtgtcccaatcgttccccgagatgtac 1680
| | | | |
Db 7446 TCAAGCTCAACCAATCCAAAGATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7505
| | | | |
QY 1681 taacatccctgaggaagtgctccttga tgaataatgaactcgtgtataacccgacaaaga 1740
| | | | |

Db 7506 TACCATCTGAGGAGAGGTGCTTTGATGAATATGACTATTCGGATATACGCCACCAAAA 7565
| | | | |
QY 1741 gaaccggtatcccttcccttaactgtatgtatgtctgagtttagggagcggtgtgctagga 1800
| | | | |
Db 7566 GAACCGGTATCCCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 7625
| | | | |
QY 1801 acagggacaagctgctcctgtatcaacagacacagcaacttagagaaagacttggtagctta 1860
| | | | |
Db 7626 ACAGGGACAGCTCCTTGATACAGAGCACCACAGCACTAGAGAAAGACTTGGTAGCTA 7685
| | | | |
QY 1861 catgcygcaatgacagaagaatctccagaccttaaaagagtcctgttagcaacctagaagag 1920
| | | | |
Db 7686 CATGCGCATGACAGAGATCTCCAGGCTTAGAGAGTCTGTAGCACTAGCAAGAG 7745
| | | | |
QY 1921 tccctgtactcttctgtctgaatgtgttctcaagaaacggagggatagaatcgtgttt 1980
| | | | |
Db 7746 TCCCTGACTTCTTGTCTGATAGGTGTTTACAGAAACCGAGGGGATTAATGATCTGCTT 7805
| | | | |
QY 1981 cttaagagaaggtgtgtatgtgtcagcttaaaagaagaatgtgtctctatgtatgaac 2040
| | | | |
Db 7806 CTTAAGAGAGGTGTGTATGTGTGAGCTTAAAGAGAAATGTGCTTCTATGTAGATCAC 7865
| | | | |
QY 2041 tcaagagccatcagagactcca tgaacaaagcttaga aaaaagtttagagagcgctgaag 2100
| | | | |
Db 7866 TCAGAGAGCATGAGAGACTCCATGAGCAAGCTTAGAGAAAGTTAGAGAGGCGTGAAGG 7925
| | | | |
QY 2101 gaaagagaagctgtgacgaaggtgtgtttgaagaaatgtgtcaacaggtctcctgtatgac 2160
| | | | |
Db 7926 GAAAGAGAGGCTGACAGGAGGTGTGTGAAGATGTTTCAACAGTCTCTTGTGATGAC 7985
| | | | |
QY 2161 accctgtctctgtctgtcgtgagggccccctagtagtccctgtcttacttaacatgtgg 2220
| | | | |
Db 7986 ACCCTGTCTTGTCTGT 8045
| | | | |
QY 2221 cctgtcttaataataggt 2280
| | | | |
Db 8046 CCTGTGTATTAATTAAGT 8105
| | | | |
QY 2281 atgtgacttagcaacagtagcaacaggtctctgacaaaggaacagcaactgaactgaact 2340
| | | | |
Db 8106 ATGTGACTTAGCAACAGTAGCAACAGGCTTGTGAGCCAAAGGAAATGACCTTGTAGCT 8165
| | | | |
QY 2341 tcccaagttcctaagatlaaactaacttaacaagaacagaaagttgagaaatgaagaa 2400
| | | | |
Db 8166 TCCAGTCTTAGATTAAGATTAAGACTATTAAACAGCAAGAGTGGGAATGAAGATGA 8225
| | | | |
QY 2401 tgaacactaaacctccagaaacccaagaaatlaataaaagctctaaatgtcccgaaat 2460
| | | | |
Db 8226 TGCACACTTAACCTTCCAGAACCCAGAAAGTTAATTAAGAAAGCTCTTAATGTCCCGAAT 8285
| | | | |
QY 2461 aagaacccctgtcgt 2520
| | | | |
Db 8286 CCAGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8345
| | | | |
QY 2521 ctgcatcctgtgcccctaaagtaagataacaggaatgaagttgactgaactgaactgaact 2580
| | | | |
Db 8346 CTGCTATCTGTGCTTAAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 8405
| | | | |
QY 2581 ctgtaaaactgtactgtgacccaatgaagaatgtatcaacatgtgacagccctagtgacta 2640
| | | | |
Db 8406 CTGTAAAACGTACGTGACCATAGAAAGTAATGATTATACATTTGACAGGCTAGTAGACTA 8465
| | | | |
QY 2641 tctcaactgaactgtgtactctgtcccaagggccccagacagatgacagctcctgagacta 2700
| | | | |
Db 8466 TCTCAACTGCAATCTGTACTCTGTCCAGAGCCACGACATGCGGACCTCCGAGACTA 8525
| | | | |
QY 2701 tttaaataatgtgtgtcagcaggaagcggtgtcctgataatttaaataatgtgtgtcat- 2759
| | | | |
Db 8526 TTTTAAATGATTTGTTCCAGGAGCGGGGCTCTCGATATTTTAAATGATTTGTTCCAGC 8585
| | | | |
QY 2760 ----- 2759
| | | | |

QY	1	tgcttttaggttagaataacccccgtagaagtttggctgtagacccctatgaattgctct	67.8%;	Score 2004.4;	DB 14;	Length 8849;
		Best Local Similarity	81.6%;	Pred. No. 0;		
		Matches 2480;	Conservative	0;	Mismatches 406;	Indels 152; Gaps
QY	61	acgggggaacccccctgttagaataatgctctgtacataagttgtagtgcgctt				
Db	5787	TGCTTTTAAggGTAGGAAACACCCCTGGAAGATTGGGCTGACCCCTAATGAATTGCTCT				5846
QY	121	ccagacctgtctctcctaagctcaagagcgctcagttgagtgaggaacagagcttgaagc				180
Db	5907	CCGAGCGCTTGTCTCTAGGCTCAAGGCGGCTGAGTGGGTGAGGCAACGAGCGTGAAGC				5966
QY	181	agctcggagagcctactcagaagagagacttgcagttccaatgcgtcccaagt				240

QY 2340 ttccagcttaagatagaactaaacaagaagatgaggaatgaagatga 2399
DB 8135 TTCCAGTTTAAAGATAGAACTATTACAGACAGAGAGTGGGAATGAAGATGA 8194
QY 2400 atgcaactaacccctccagaagccagaagatlaataaagcttaagtc 2459
DB 8195 ATGCAACTAACCCCTCCAGAACCCAGAGAAATTAATAAAGCTTAATCC 8254
QY 2460 tacagaacctgctgctccagtaaatgagtagaagatgcaactctatgt 2519
DB 8255 TCCAGACCCCTGCTGGCTCCAGTAATAGTAGAAGGTCACACTTCTATGTTCC 8314
QY 2520 cctgctatcctggcctaaagtaagaataaagaaatgagtaactatgct 2579
DB 8315 CCTGCTATCTCGGCTTAAGTAAGTAACAGGAATGAGTTGACTTAATCGCT 8374
QY 2580 tcttaaaactgactgagacataagaagatgattacacttgacagccctag 2639
DB 8375 TCTGTAACACTGACTGGCACCATTGAAGAAATGATTACACTGACAGCCCT 8434
QY 2640 atctcaactgcaatctgcaactctgcccagaagccagcagatgagacc 2698
DB 8435 ATCTCAACTGCAATCTGTCACTTGCCAGAGCCCAAGCAGATGCGGACT 8494
QY 2699 -----tatttaaaatgattgctcac 2720
DB 8495 ATTTTAAATGATTGTTCCACGAGCGCGGCTCTCGATATTAAATGATTG 8554
QY 2721 ggaagcgggagctcgcgaatttaaatgattgctcagagagcgagctc 2780
DB 8555 GGAAGCGGCTCTCGATATTAAATGATTGTTCAATGAGCGCGGCTCTCGAT 8614
QY 2781 ttaaaatgattgctgagcagcagagcttgctggaacccataaagctg 2840
DB 8615 TTAATGATTGTTGTTGAGCAGCAGGCTTGTGTAAGCCCAATAAAGCTTCC 8674
QY 2841 ttccgcaactgggagcgagactcctacccctgctgctgagcagctg 2900
DB 8675 TTCCGCACTCGGCGCCAGCTCTACCCCTGCTGTTGACGACTGCGGCC 8734
QY 2901 cgcttgaataaaatcctctgctgctgctgctgctgctgctgctg 2938
DB 8735 CGCTTGAATATAAATCTCTGCTGTTGCTGATCAAGA 8772

RESULT 10
A66559
LOCUS A66559 1974 bp DNA PAT 29-MAR-1999
DEFINITION Sequence 9 from Patent WO9740167.
ACCESSION A66559
VERSION A66559.1 GI:4538112
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.

REFERENCE
1 (bases 1 to 1974)
AUTHORS Galbraith,D.N., Haworth,C., Lees,G.M. and Smith,K.T.
TITLE PORCINE RETROVIRUS
JOURNAL Patent: WO 9740167-A 9 30-OCT-1997;
Q ONE BIOTECH LTD (GB)
Location/Qualifiers
Source 1. 1974
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 534 a 489 c 483 g 468 t
ORIGIN

Query Match 66.7%; Score 1970.8; DB 6; Length 1974;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1972; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 364 atgcatcccaagcttaagctgagcaccctccgactcgaggatgagagccgcaaaagactg 423

DB 1 ATGATCCCAAGTTAAGCTGGCGCAACCTCCCACTGCGGGTGGAGCGCAAAACAG 60
QY 424 agaatcccttaagcttgccctcaatgcgctggttcccttaacttaacaactcccg 483
DB 61 AGAATCCCTTAAGCTTGCGCTCCATCGCTGCTCTTCACTTAACAATAACTCC 120
QY 484 gccagtagtaaacgcttataagacagctgaaccccatagaccttataccctactcg 543
DB 121 GCCAGTAGTAAGCGCTTATAGACAGCTCCAAACCCCATAGACCTTATCCCTAC 180
QY 544 ctgattatgacccctgatacagggtgctgactgtaaalagacactgaggtgtgctc 603
DB 181 CTGATTATTAGCCCTGATACGGGTGACACTGTAATAGCACTGAGAGTGTGCTCT 240
QY 604 ggcacctgggtgacctgaactgcaactctgctcgcctcagcttgattaaacccg 663
DB 241 GGCACCTGGTGGCTTAACCTGCAATTTCTGCTCCGATGATTAACCCCTGTTAA 300
QY 664 acacccccaactagctcgaattgagttcattgctgcccagagcaagaaagag 723
DB 301 ACACCTGCCAAGCTAGTCCCTAGTTATGGTTCTATTGCTGCCAGGACAGAGAA 360
QY 724 aaatgctgagggttctggggaatcctctgtaagagatgagctgctgcaactc 783
DB 361 AAATAGTGTGGGCTTCTGGGAATCTCTGTAGAGATGAGAGTGGCTCACTCAAC 420
QY 784 gatgagagctggaatgagcagatctctccagagccggtaaatctcccttgcaat 843
DB 421 GATGAGACTGGAAATGGCGGATCTCTCCAGAGCCGGGTAAATTTCTTTGTCA 480
QY 844 tccgcccggcgaagtaacaatgataagaaactataaagaataagagctgctcc 903
DB 481 TCCGGCCCGGCAAGTACAAATGATGAATCTATTAAGATAGAGCTGCTCCCAT 540
QY 904 gacttgattatcctaagaatgattcactgaaagagaaagaaatlaacaaag 963
DB 541 GACTTGATTATCTAAAGATTAAGTTCACTGAAAAAGCAAAACAGAAATATTCA 600
QY 964 tggataatgattgagctggggaatgatttlaaataatgagggggaagaggtgc 1023
DB 601 TGGATTAATGATTGAGCTGGGGAAATGTTTTTATAAATATGCGGGGACAGG 660
QY 1024 accttaacatcgccttgagatagagcgggagcaaacccctgtgcaatggaacc 1083
DB 661 ACTTTAACATTCGCTTAGATAGAGAGGAGGACGAACCCCTGTGCAATGGAG 720
QY 1084 gataaagtaactgctgaaagagggcccgccctgagagcagcagataactgtc 1143
DB 721 GATTAAGTACTGCTGAACAGGGGCCCGGCTTGAGCCACCGCATTAATTCG 780
QY 1144 ccccaatlaacctgctgctgagctgacataaacaagcgcctagcaacagatcac 1203
DB 781 CCCCAATTAACCTGCTGGGCTTGACATTAACACGCCCTGAGCAACGTCCACT 840
QY 1204 ttgattcccaacaacgagcctagaactcccaaggtgttcctggttaagaacaga 1263
DB 841 TTGATTCTTACCAACAGCGCTAGAAATCTCCAGGTTCTCTGTTAAGCAGAG 900
QY 1264 ctcttgaatcctatccaagggagcttccaagccatcaactccaacgacctga 1323
DB 901 CTTCTAGTTCATCCAGGAGCTTCCAAAGCATCACTCCACCGACCTGATCC 960
QY 1324 tctctgtgtgcttgtctatcccaagggcctccttataatgagggatgctaa 1383
DB 961 TCTTCTTGTGGCTTTGTTATCTCAGGCGCTCTTATTAGAGGAGTGGCTTAA 1020
QY 1384 agaaaaatcaatgtaaccaaaagatagaatgaatgaatgaggggtccgaaata 1443
DB 1021 AGAAAAATTAAATGTGACCAAAAGAGATGAATTAATGTACATGGGGGTCCGAA 1080
QY 1444 cttaacctactgaagtttccgggaaggagatgacatagaaagctcccccatacc 1503

Db 1081 CTACCTCAGTGAAGTTCCGGAGGAGGACATGATAGAGAAAGCTCCCAATCCAC 1140
 QY 1504 caaaccttgcctagtagtctgttctagagcagcctcaagaatactagattagta 1563
 Db 1141 CAACACCTTGTCTAGTACTGCTGTTATGAGAGGCGCTCAGAAATCGATTTAGTA 1200
 QY 1564 cctgttatacaagtggtgagcagtaactggtgttccccctgtgttccacctca 1623
 Db 1201 CCTGGTATATACAGGTGGTGGCATGCAATCTGGTTAACCCCTGTGTTCCACCTCA 1260
 QY 1644 gtcttcaaccaatcaagaattgtgtctatggttccaaatgtccccagtgctactac 1683
 Db 1261 GTCTTCAACCAATCCAAAGATTGTGTCTCATGGTCCAAATCGTCCCGAGTGTACTAC 1320
 QY 1684 catcctaagaagtggtcctctgataatgatactatgataatcagcagcagcaagaaga 1743
 Db 1321 CATCTGAGGAGAGTGTGCTTGTGATGATGATGATGATGATGATGATGATGATGATG 1380
 QY 1744 cccgtatcccttaccctagctgttaatgctcgattaggaagcgcgttggcgttagaaga 1803
 Db 1381 CCCGTATCCCTTACCTAGCTGTGATGCTGCGATTAGGAGCGCCGCTGTGAGAGACA 1440
 QY 1804 gggagacgtccctgatacaagaagcagcagcagcagcagcagcagcagcagcagcagc 1863
 Db 1441 GGGAGAGTGCCTGATCAGAGGACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1500
 QY 1864 ggcgcacatgaagaagatctccagcagcagcagcagcagcagcagcagcagcagcagc 1923
 Db 1501 GCGGCATGACAGAGATCTCCGAGCCTTAAAGAGATCTGTAGCAACCTAGAGAGATCC 1560
 QY 1924 ctgactcttctgtctgaagtgtgtctcagaaacgcgagggagattagatctgcgttctca 1983
 Db 1561 CTGACTTCTTGTCTGAGAGTGTCTGACAGAACCGAGGGGATTTAGATCTGCTTTCTA 1620
 QY 1984 agaaagtggtgtatgtcagcagcagcagcagcagcagcagcagcagcagcagcagc 2043
 Db 1621 AGAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680
 QY 2044 ggaagcactcagagatccatgatacaagcagcagcagcagcagcagcagcagcagcag 2103
 Db 1681 GGAGCCATCAAGAGACTCCATGAAACAGCTTGAAGAAAGTTAGAGAGCGCTGCAAGGGA 1740
 QY 2104 agaaagctgagcaggggtgtgttgaagaagtgttcaacagcagcagcagcagcagcagc 2163
 Db 1741 AGAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1800
 QY 2164 ctgcttctcgtctcagcagggccccctagtagctcgtcgtctgttacttaagtgagcct 2223
 Db 1801 CTGCTTCTGCTCTGAGAGGGGGCCCCCTAGTAGTCTGCTCTGCTTACTTACAGTTGGGCT 1860
 QY 2224 tgccttaataagtggttctgtccttctgttgaagaacagtgagtcagtcagatcatg 2283
 Db 1861 TGCCTTAATTAATAGTGTGTGCTTGTGAGAAACAGAGTGTGAGTGTGAGTGTGAGTGT 1920
 QY 2284 gtacttgagcagagtaaccaagcctctcagcagcagcagcagcagcagcagcagcagc 2337
 Db 1921 GTACTTATGAGCAGAGTGTGAGAGGCTTCTGAGCCAGGAAAGTACTGCTCTAG 1974

RESULT 11
 AF014162
 LOCUS AF014162 1971 bp mRNA VRL 01-JUN-1998
 DEFINITION Porcine endogenous retrovirus envelope protein (env) mRNA, complete cds.
 ACCESSION AF014162
 VERSION AF014162.1 GI:3169735
 KEYWORDS
 SOURCE porcine endogenous retrovirus.
 ORGANISM
 Viruses: Retrovirus; Retroviridae; Mammalian type C
 retroviruses; 1-Mammalian type C virus group.
 REFERENCE
 1 (bases 1 to 1971)
 AUTHORS Hamworth, C., Galbraith, D.N., Lees, G.M. and Smith, K.T.

TITLE Porcine endogenous retrovirus (POEV) Env sequence
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1971)
 AUTHORS Hamworth, C., Galbraith, D.N., Lees, G.M. and Smith, K.T.
 TITLE Direct Submission
 JOURNAL Submitted (16-JUL-1997) Q-One Biotech Ltd., Todd Campus, Acre Rd., Glasgow G20 0XA, UK
 FEATURES
 source
 1..1971
 /organism="porcine endogenous retrovirus"
 /db_xref="taxon:61673"
 /cell_line="derived from cell line PK15"
 /note="PoEV"
 1..1971
 /gene="env"
 1..1971
 /gene="env"
 /codon_start=1
 /product="envelope protein"
 /protein_id="AAC17948.1"
 /db_xref="GI:3169736"
 /translation="MHPTLSRRHLPTRGEPKRLRILPFSASIMFLTLTPPOASSK
 RLINSHRPLSLTLWLIIDDTGVTVNSTRGVAPRGWMPBELHRLINPAKSTP
 PNLVNSYFPCPEPEKREKKGSGESGECRMSCTVSTNDGDKRISLQRYKSPN
 SGPGRYKMKLTKDSCSDLDYLTSTFKRTGKTSKDKWT ELDGNSFLYEGGAG
 STLTIRLTETGEPVAMGPDVLAELGEPALPEPNNLPVPLTSLRPTDPPGNS
 TGTILRPTPTNSPBPVKYTORLEFSLFOAIINSTDPATSSCMJCSGPPYE
 GMAKERNVTRKHNOCITGSRNKLITELVSGKGTGIGKAPSHOHLCTVYEDG
 SENORLVGYRHWACNTGLAPCVSTVFNSKFCVWVOIPEVYHPEVYDEVD
 YRNPFRPEVSTLVAVLIGTAVGAGTGTGALITRPOLEKGLGLHMAFEDLRA
 LKESVSNLESLSLSEVYLDNRGIDLLPLRREGILCAALKRECCFVYDHSGARDSM
 NKLKRLRRRRRADDGMEFEGMNSPMWTLTSLALGVLVLLVLLVLPCLINK
 FVAFFREVSAVOIIMVLRKQYOGLSGSEIDL"

BASE COUNT 530 a 490 c 482 g 469 t
 ORIGIN

Query Match 65.7%; Score 1941.6; DB 14; Length 1971;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 1967; Conservative 0; Mismatches 4; Indels 3; Gaps 2;

QY 364 atgacccacagcttaagctgagcagcagcagcagcagcagcagcagcagcagcagcagc 423
 Db 1 ATGCATCCCACTTAAAGCGGGCCCACTCCGCACTCGGGGTGGAGAGCCGAAAGACTG 60
 QY 424 agaatcccttaagctcgcctccatcgcctgcttctactactaacaataactccag 483
 Db 61 AGAATCCCTTAAAGCTTGCCTCCCAAGCGCTGCTTCTTACTTAACAATAACTCCAG 120
 QY 484 gccagtagtaaacgcttataagacgctcgaaccccccatagaccttaccctaccg 543
 Db 121 GCCAGTAGTAAAGCGCTTATAGACAGCTGCAACCCCAATAGACCTTTATCCCTTACCTGG 180
 QY 544 ctgattatgacccatgatacgggtgctcactgtaataagcaacgagcagcagcagcagc 603
 Db 181 CTGATTATGACCCCATATGCGGGGTGCTCACTGTAATAGACACCGAGGTGCTCTTGA 240
 QY 604 ggcacctggtgacctgaactgacttctgctcgcagattgataaccgcgctgttaaacg 663
 Db 241 GGCACCTGTGTGCGCTGAACTGATGATGATGATGATGATGATGATGATGATGATGATG 300
 QY 664 aaaccccccaactgctcgtatgattgattgattgattgattgattgattgattgattgatt 723
 Db 301 ACACCTCCCAACCTATGCTCGTATGAGGTTCTATGCTGCCAGGACAGAGAAAGAG 360
 QY 724 aaatactgtggggtgtcttgaggaatcctctctgagagatgagcagcagcagcagc 783
 Db 361 AATAATGTGGGGGTCTTGGGAATCTCTCTAGAGATGAGAGTGTGCTCACTCCCAAC 420
 QY 784 gatgagagctggaatgagcagatctctcagagacgggtgtaaaatctccttctgcaat 843
 Db 421 GATGAGAGCTGGAATGAGGCGGATCTCTCCAGAGCGGGTAAATCTCTTGTCTCAAT 480

QY	844	tcgcgcccgggaaagatcaaaaaatgatgaaacatctaaaga-aaagagctgtccccaatca	903
Db	481	TCGCGCCCCGGGCAAGTACAAATATGATGAACATATATAAGATAAAGACTGCTCCCATCA	540
QY	904	gaactagattatctaaagaataagttctactgtgaaaaaggaacagaaatatctccaaag	963
Db	541	GACTGTAGATTATCTTAAAGATAAGTTTCACTG--AAAGGAAACAGGAAATATTCAAAG	598
QY	964	tggataaaygtatlgagcttggggaaatagttttttaaataatgycggggagcaagylcc	1023
Db	599	TGGATTAATCGTATGAGCTGGGGAAATAGTTTTTT-ATTATATGGGGGGAGCAGGGTCC	657
QY	1024	actttaaccaatctgccttggataagagaaagggaaacaaacccctgtgtcaatlgjaccc	1083
Db	658	ACCTTAACCAATTCGCTTAGGATATGAGAGGGGACAGAACCCCTGTGTGCAATGGGACC	717
QY	1084	gataaaatlaactggtctaaacaggggccccggccctggagccacgcgcataacttgcgggt	1143
Db	718	GATTAATTAATGTGGCTTAACAGSGGGCCCCGGCCCTGTGAGACCAGCATTAACCTTGGCGTG	777
QY	1144	ccccaaatlaacctgcgtgcggcctgtgacataaacagccgcctagcaacagtacaacttga	1203
Db	778	CCCCAATTAACCTGCCTGGCGGCTGTGCATTAACACAGCCGCTTACAGCAATACACTGGA	837
QY	1204	tgtattcttaaccaaacgcgtctgaagaactccccaagtlgtccctgttaagaagcggcgaaga	1263
Db	838	TTGATTTCTTAACCAACACGCCCTAGAAACTCCCAAGGTGTTCTGTTTAAGACAGGACAGGA	897
QY	1264	ctcttaagctcaatccagggaggttccaaagccataacttaaccgcagccctgtatgcact	1323
Db	898	CTCTCAAGTCTCATCCAGSAGCTTTCANAGCATCAATCCAGCGACCCCTCATGCCACT	957
QY	1324	tctctgtltgtgcltltgctatccctcaaggccctccattatlgaggatgtgctaaagaa	1383
Db	958	TCTTCTGTGGCTTGTGTCTATCTCAGGGCCCTCTTATTATGAGGAGATGGCTAAAGAA	1017
Db	1018	AGAAATATTCATGTGACCAAGAGCAATAAATATCAATTATCATGATGGGGGTCCGGAATTAAG	1077
QY	1444	cttaacctcaactaaagtttccggaaagggagactgcatatggaaaagctcccccataccac	1503
Db	1078	CTTAACCTCACTGAAGTTTCCGGGAAGGGGACATGATATGAAAAAGTCCCATTCACAC	1137
QY	1504	caaacaccttgcatagtactgtgttltatlgacagggccctcgaagaataatagattagta	1563
Db	1138	CAACACCTTGTGATATGTAAGTATGHTTATMGACAGACGCTCAGAAATACATTAATTAGTA	1197
QY	1564	ccgtgttatataacaggtgtgtggcacaatcacttgggttaacccctgtgttctcaactca	1623
Db	1198	CTGTGTATTAACAGGTGGTGGGATCAATACAGGGGTTTAACCCCTGTGTCTCCACCTCA	1257
QY	1624	gtcttcaacaatccaaaagtttgtgtgtcatgtgtccaaatgtctcccgagttgtaactac	1683
Db	1258	GTTCTTCAACCAATCCAAAGATTTCTGTGTATGTGTCCAATGTCTCCCGAGTGTACTAC	1317
QY	1684	catctcttgaagaaatgtgttcctcttgaatgaatlgactatcgtlataacgcacaaagaagaa	1743
Db	1318	CATCTGTAGGAAGTGGTCTTGATGAATATGACTATCGGTATTAACGACCAAAAAGACAA	1377
QY	1744	cccgatcccttaacctagctgtatcatgtctcogattagggagcggcttggcgttagaaga	1803
Db	1378	CCCCATTAACCTTACCCCTAGCTGTAAATGCTTCGCGATTATAGGACGCGGTTGGGCTTAAAGAA	1437
QY	1804	ggagcagatccctctgatcatcagaagcaacagcaactcagaagaagagcttgttagctatc	1866
Db	1438	GGGACAGCTGCCCTGATACAGACACCAACAGCACTGTGAATAAAGAGACTTGTTAGCTTAAT	1497
QY	1864	ggcgacatlgacagaagaatctccgagcctttaaagaaatctgttagcaacctagaagaatcc	1923
Db	1498	GGCGCCATGACAGAAAGATCTCCGAGCCTTAAAGAGTCTGTATAGCAACTTAAGACAGTCC	1557
QY	1924	ctggaactcttgtctgaagatgtgttctcaagaaacgggaaggtatagatctgtcttcta	1987

Db	1558	CTGACTTCTTTGCTGTAAGTGGTTCTACAGAACCGAGGGGATAGATCTGTTTCTTA	1617
QY	1984	agagaaggttgggttattatgttgcagcctttaaagaagaatgttgcctctatgtatgcactta	2043
Db	1618	AGAAAGAGTGGCTATGTGTCAGCCCTTAAAGAGAGATGTTCTCTTATGTAGATCACTTA	1677
QY	2044	ggaagccatcagagactcccatgtacaacaactcttgaaaaaagttagagagagcttgaagagaa	2103
Db	1678	GGAGCCATCGAGACACTCATGTGAACAACCTTAGAAAAAAGTTAGAGAGCGCTCCAGGGAA	1737
QY	2104	agagaagcttgaccagaggttggcttgaagatagtttcaacaggtctcctttgatggaccacc	2163
Db	1738	AGAGAGCGCTCACCGAGGGGGTGGTTGAAGATGCTTCAACAGCTCCCTTGGAATGACCACC	1797
QY	2164	cttgcttttgccttgtagcaggggcccccttaagtgcctgcctcctgttacttaeaagtttgacct	2223
Db	1798	CTGCTTTTCTGCTGACGGGGGCCCTCAGTAGCTCGCTCGTTACTTAAAGTTGGGCCCT	1857
QY	2224	tgacttaataataggttcttgcccttcttgaagaacagtgatgtgcacgtccagatcatatg	2283
Db	1858	TGCTTAATTAATATAGGTTTGTTCCTTTGTTACAGACACAGATGCAATGCCATGCATCATG	1917
QY	2284	gtacttggcaaacagtaaccaagcctcttgagccaagagaagaactgtaccctctag	2337
Db	1918	GTACTTATGGCAACAGTACCAGGCCCTTTCGAGCCCAAGAGAAAGTAAGTGCACCTCTAG	1971

RESULT	12
LOCUS	AF038601
DEFINITION	AF038601 7333 bp mRNA MAM 21-MAY-1998
ACCESSION	AF038601
VERSION	AF038601
KEYWORDS	Sequence.
SOURCE	AF038601.1 GI:3133304
ORGANISM	pig.
REFERENCE	Sus scrofa
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
TITLE	1 (bases 1 to 7333)
JOURNAL	Akiyoshi,D.E., Denaro,M., Zhu,H., Greenstein,J.L., Banerjee,P. and
MEDLINE	Fishman,J.A.
REFERENCE	Identification of a full-length cDNA for an endogenous retrovirus
AUTHORS	of miniature swine
TITLE	J. Virol. 72 (5), 4503-4507 (1998)
JOURNAL	2 (bases 1 to 7333)
MEDLINE	98216827
REFERENCE	Fishman,J.A.
AUTHORS	Direct Submission
TITLE	Submitted (16-DEC-1997) Infectious Disease Unit, Massachusetts
JOURNAL	General Hospital, 55 Fruit Street, Boston, MA 02114, USA
MEDLINE	3 (bases 1 to 7333)
REFERENCE	Fishman,J.A.
AUTHORS	Direct Submission
TITLE	Submitted (15-MAY-1998) Infectious Disease Unit, Massachusetts
JOURNAL	General Hospital, 55 Fruit Street, Boston, MA 02114, USA
MEDLINE	Sequence update by submitter
REFERENCE	On May 16, 1998 this sequence version replaced gi:3116445.
COMMENT	Location/Qualifiers
FEATURES	1..7333
SOURCE	/organism="Sus scrofa"
	/strain="miniature swine"
	/db_xref="taxon:9823"
	/cell_line="PK-15"
	/note="porcine endogenous retrovirus ERV-PK15; infective
	virus but represents a naturally occurring transcript"
	1..578
LTR	598..2172
CDS	/codon_start=1
	/product="gag protein"
	/protein_id="AAC16766.1"
	/db_xref="GI:3116446"

```
/translation="MGQVTPPLSLIDHTEVSRHNLISVYKKKPPWOTFCASENP
FDVWPSSEGTSEIILAVNAIILFQPGSHPOQEPYIITWQDLADPPWYKPMWLN
KPRKGPRIILGEMKNSHAKVEPSRIYPIEIEPTWEPYPPPYAPWCAVGS
PSAPGAPVWEPAGTRSRGATPERTDEIAIILPLRYGPMGGOLOIOWYFSS
ADLYNMKINHPFESDPORLTGLVLSMFESHQPMCCOLOLITFTTEERITLLEA
KKVNGADGPTOLONEIDMGEPILTRMGVNTAEGRSEKIKYAOAVAGIRGSRP
TNIAKRVWMOGPMPPSVIERLEMAFRRTPTDPDSEAKASVALATIGOSALDR
KKIOLREIOLAEIIRDLVREAEKVIYFREIEEKEKREEREREERERORREKNT
KILAAVEGSSSRREDFRIRSGPROSGLNRPILDKDOCAVCEKKGHWANMCPK
KNGKPVILAEEDK"
CDS
2320..3552
/contig="truncated protein"
/codon_start=1
/product="pol protein"
/protein_id="AAC16767.1"
/db_xref="Gen:3133305"
/translation="MGATGORYPMTTRPVLDGVHVSFLVIEPCPVLLGRDL
TKMGAQISFEGREPSVNNKPITVLTLQDDEKRLSPVKKPRODIOSLBEPQAM
AETAGMGLAKOVPROVIOUKASATPVSRVYPLSRREBSIIMHYQLIDGILLVPO
SPWMTPLIPRKRETDNDYRYQDLREYNKRVQDIHPVPMYNDLSALPERMWTVL
DKDAFFCLRHPSTQPLFTFEMWDPGTGOTWTRLPQGNSTPIDELHRL
ANFRIOHPVYTLLOYVDLLAGATKODCEGRKALILEISDLGYSRAKKAQICRE
VTVLGYSLRGGLWLEARKKTVQVQIPAPTTAKOVREPLGTAGRCRLMIPFATLAP
LYPLTKKGGCLPQGGK"
4803..6725
/contig="env protein; N terminus deleted"
/codon_start=1
/pseudo
LTR
BASE COUNT 1984 a 1788 c 1885 g 1676 t
ORIGIN
Query Match 63.7%; Score 1882.8; DB 4; Length 7333;
Best Local Similarity 80.1%; Pred. No. 0;
Matches 2409; Conservative 0; Mismatches 402; Indels 195; Gaps 9;
CDS
1 tgccttttaggttagaacaacccctggaagcttggcctggaacccctatgaattgctct 60
4462 TGCCTTTTAGGTTAGGAGAACCCCTGGACAGTTGGCTGACCCCTATMAATTGCTCT 4521
QY 61 acggggagaccccccccgctgtgtagaattgctctgtagaagctgtagaattgctct 120
Db 4462 TGCCTTTTAGGTTAGGAGAACCCCTGGACAGTTGGCTGACCCCTATMAATTGCTCT 4521
QY 4522 ACGGGGAGACCCCGCTGGACAGAAATGCTTGCATAGTGTGATGCTGCTT 4581
Db 4522 ACGGGGAGACCCCGCTGGACAGAAATGCTTGCATAGTGTGATGCTGCTT 4581
QY 121 cccgacctgttctcttagagctcaaggcgtcgagtgagtgaggaacagcgctggaac 180
Db 4582 CCCAGCCTTTGTCTCTAGCTCAAGGCGCTCGAGTGGGTGAGGACGACGCGTGAAGC 4641
QY 181 agcccgaggagcctactcaggaagagagagcttgaagttccaatcgtcttccagt 240
Db 4642 AGCTCCGGAGGCTACTCAAGAG---GAGACTTGAAGTTCACATCGCTTCCAAATTG 4698
QY 241 gaattcaagtctatagtagagcagcgttgagaagaaccccggaacccgctggaagagac 300
Db 4699 GAGATTCAAGTCTATAGACGCCACCTGTGAGAAACCTCGAGACTCGGTGAGAGGAC 4758
QY 301 ctatctcgtacttttagaccacacacagcgctggaagtcgaagagatcctccacgtga 360
Db 4759 CTATCTCGTACTTTTGGACACACCAACGGGCTGTGAAGTCGAAG-----4803
QY 361 tcatatcatcccaagtttaagctgagccacctccgaactcggggttgagagacgaaaaa 420
Db 4804 -----4803
QY 421 ctgagaatcccttaagctcgtcatcgtcgtgtctcttactactaaataactcc 480
Db 4804 -----GAATCCCTTAAGCTTGGCTCCATCGCGTGTCTCTTACTGTGTAATAACTCTCT 4859
QY 481 caagccagtagtaacgacctatagacagctgaaccccatagaccttatccctacc 540
Db 4860 CAACTTAATGTAAGCCCTTGTGGACAGCCGCACTCCCATTAACCCCTTATCTCTCAC 4919
```


Db 4804 ----- 4803
Qy 421 ctgagaatcccttaagcttcgctccatcgctggttcctactactaacaatactccc 480
Db 4804 ----GAATCCCGTTAAGCTTGGCTCCGATCGGCTGTTCTTACTCTGTCAATTAACCTCT 4859
Qy 481 caggccagtaagtaaacgctctatagacagctcgaaaccccatagacattatccctaac 540
Db 4860 CAAGTTAATGTAAGCGCTTGTGAGCAAGCCCACTCCCAATTAACCTTATCTCTCAAC 4919
Qy 541 tggctgattattgaccctgatacgggtgtcaactgtatgaataagacaggggtgtctct 600
Db 4920 TGGTTACTTACTGACTCCGGTACAGGTATTAATTAACAGCACTCAAGGGGAGGCTCC 4979
Qy 601 aagagacccctgtgctgacactgcatcttcgctccgactgattatcaaccgctgtta-- 658
Db 4980 TTGGGGACCTGGTGGCTGTAATTAATGTCCTGTCATGATGATATCCGCTGCTCAAT 5039
Qy 659 ----aagacacccctcccaactagtcgtagttatggttctctatgtctcccaag---c 711
Db 5040 GACCAAGGCCACACCCCGATGCTACGCTGCTTACGGGTTTACGTTTGGCCAGAGACC 5099
Qy 712 acagagaaagagaaatactgtggtggtctctgaggaactcctctgtaggaagatgagctgc 771
Db 5100 CCAATTAATGAAGATATTGTGGAATCTCTCAGATTCTTTTGCACCAATGGAAGCTGC 5159
Qy 712 gtaacctcaacagatgagactggaatgacgactctctccagacgggttaaatc 831
Db 5160 ATAATCTTAATGATGGGAATGTGAATGGCAGCTCTCAGCAAGACAGATGAATGATAC 5219
Qy 832 tccctgtg-----caa 842
Db 5220 TCTTTTGTAAACATCTACAGTTAATCAATTAATTAATGCGCATGGAGATGAGAA 5279
Qy 843 ttccggccgggcaagtaacaaatgataaactatataagaataagactctcccatc 902
Db 5280 GATTGGCAACACGGGTACAAAAGATG-TACGAATTAAGCAATTAAGCTCTCAATTCCT 5338
Qy 903 agacttagatatactaaagataaagtttcaactgaaagaaagaaagaaatattcaaaa 962
Db 5339 AGACTAGATTACTTAATAAATAAGTTTCATGAAAAAGAAACAGAAATATTTCAAA 5398
Qy 963 gtgataaagtgtatgagctggtggaatagttttta---ataatgctggggagag 1019
Db 5399 GTGGTAATGTAATGTAATCTTGGGAATAAGTACTATGAGAGCTCTGAGAAAGAAAG 5458
Qy 1020 gtccacttaaccattgctccttagatagagagcggaacagacccctgtggaatgg 1079
Db 5459 ATCTCTTGTGACTATTCGCTCAGAAATTAAGAACTCAGATGAGACCTCGGTTGCTATAGG 5518
Qy 1080 acccgataaagtaactgctgacagagggcccgccctggaagcagacgcataactgac 1139
Db 5519 ACCAAATTAAGGGTTGGCCGAACAAGACCTCAATCOAGA----- 5560
Qy 1140 ggtgcccccaataactcgtcgtggtgctgatacaacagccgctgacacagtlacac 1199
Db 5561 -----ACAGAGGCCATCTCTTAACCCCTCTATTAATTAATCAATCAACCTC 5602
Qy 1200 tggatgtatctctacaaacagcctagaacatcccccaggtgttcctgtaagcagaca 1259
Db 5603 TGGATCACTCCCACTGA-----GCTTAACATCACTTAATTAACAGGGGC 5647
Qy 1260 gagactctcagtlctacagagagcttccaaagcatcaactccacgacccgtgtgc 1319
Db 5648 GAAACTTTTATGCTCATCCAGGAGCTTTCAAGCTTTAATCAGCCAGACTCCAGAGGC 5707
Qy 1320 caactctctgtgtgctgtgctatccctgaagcctccttataatgaagggatggtctaa 1379
Db 5708 TACCTCTCTTGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 5767
Qy 1380 agaaagaaatcaatgtgacaaagacatagaatcaatgtacatgggggtcccgaaa 1439
Db 5768 AGGAGGAATTAATGATGACAAAGAAACATAGAACCAATTAATGATGGGATCCCAAA 5827
Qy 1440 taagctaacctcaactgaagttccgggaagggagacatgataagaaactcccatc 1499
Db 5828 TAAGCTTAACCTTACTGAGGTTTCTGGAAGAAAGCACCCTGATAGGATGGTTCCCATC 5887
Qy 1500 ccacaaacaccttctgctatagtaactgtgttctatgagcaagcctcagaaatcagtaatt 1559
Db 5888 CCACCAACACCTTGTGAACACACTGAAGCCTTAATGAACCTGTGAGTCAATATATCT 5947
Qy 1560 agtagctgttataacaggt 1619
Db 5948 GGTACTGTTATGATGAGGT 6007
Qy 1620 ctcaactcaacccaatcccaagattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1679
Db 6008 CTGTTGTTTCAACCAACTAAAGCTTTTTCGTTATGTTGTTCAATATGTCCCGGGGTGA 6067
Qy 1680 ctacacatcctgaggaagt 1739
Db 6068 CTACTATCCGAAAGAACAGCTCTGATGATATGACTATATATATATATATATATATATAT 6127
Qy 1740 aagaacccgtatcccttaacccctgaactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1799
Db 6128 AGAGCCCATATCCCTGACACTACTGTATATGATGATGATGATGATGATGATGATGATGAT 6187
Qy 1800 aacagagacagctgacccctgaacagagacacagcagctagagaaagagactgtgtgact 1859
Db 6188 AACGAGAGAGGCTGCCCTATATCAAGAGACCGCAACAGCTGTGAAGAAAGACTTATGATCT 6247
Qy 1860 acatgagcaatgacagaaatctcgaagccttaagagtgctgttagaacactagaaga 1919
Db 6248 ACATGCAATGTAAAGCAAGATCTCAAGCCCTAGAAAATCTGTCAATACCTGAGAGA 6307
Qy 1920 gtccctgactcttctgt 1979
Db 6308 ATCCCTAACCTCTTATCTGAAAGTGTCTTACGAAACAAAGGGGTTGATCTGTTAT 6367
Qy 1980 tctaaagaaaggt 2039
Db 6368 TCTAAAGAAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6427
Qy 2040 ctcaagagacatgaagactcctgaacaaagcttgaagaaagcttagagagcgctcgaag 2099
Db 6428 CTCAGAGCCATCAGAGACTCCATGAGCAAGCTTAGAGAAAGGTTAGAGAGCGTGAAG 6487
Qy 2100 ggaagagagaggt 2159
Db 6488 GGAAGAGAGGCTGACCAAGGGGTGTGAAGGATGTTCAACAGGCTCTTGGATGAC 6547
Qy 2160 caccctgtcttctgctcctgaagggcccttagtagtccctgctcctgttactacagtgg 2219
Db 6548 CACCGTCTTTCGCTGCTAGAGGGCCCTAGTAGCTCTGCTCTTACTTACAGTTGG 6607
Qy 2220 gectgttaataatagttgt 2279
Db 6608 GCTTGTCTTAATTAATAGTTGT 6667
Qy 2280 catgtacttaggcaacatlaccaaggtctctgagccaaggaagaaactgactctagcc 2339
Db 6668 CATGTACTTAGGCAACAGTACCAAGCCTTGTGAGCCCAAGAGAAAGTGAACCTGTATACC 6727
Qy 2340 ttccagttctaaagattagaactataaacaagaagaagtggtggaatgaagaatgaa 2399
Db 6728 TTCCAGTTCTTAAGATTAACTATTAAACAACAAGAGTGGGAATGAAGATGA 6787
Qy 2400 atgcaacctaaccctcccaagacccaagaagttaaataaagcttcaatcccccgaat 2459
Db 6788 ATGCAACCTAACCTCCCGAAGCCAGGAAGTTAATAAAGCTTAAATATGCCCCGAT 6847
Qy 2460 tacagacccctgt 2519
Db 6848 TCCAGACCTGT 6907

QY 2520 cctgctatcctgagcctaagaataacaggaatgagtgacttaactcgtatcctgagat 2579
|||||
Db 6908 CCTGCTATMCTGGCCCTAAGTAAAGATACAGGAATAGTGGTTCATCGTATTCGAT 6967
QY 2380 tctgtaactgactgagccatagaagaattgattacacattgacagccctagtacct 2639
|||||
Db 6968 TCTGTAAACCTGACTGGCACCATTGAAAGAAATTGATTACACATTGACAGCCCTAGTGACCT 7027
QY 2640 atcccaactgaactgctgactctgcccagaagcccaagatgagagactccgaagct 2699
|||||
Db 7028 ATCTCACTGCAATCTCTCCTCTGCCCCAGAGCCACAGCAATGCGGACTCCGGAGCT 7087
QY 2700 atttaaaatgattgtgtccacagagcgagcgtctcatalatttaaaatgattgtccat 2759
|||||
Db 7088 ATTATAAATGATTGGTCCACAGGAGCGGCGCTCTCATTTTAAATGATTGTGTCAT 7147
QY 2760 ggaagcgagcgtctcgatatttaaaatgattgtgtgagcgacagagcctgtgtga 2819
|||||
Db 7148 GGACCGCGGCTCTCGATATTATAAATGATTGTGTGACGACAGCGCTTGTGTGA 7207
QY 2820 accccaataaagctgtctccgactccgactcgagcgagcgtctcactacccctgctgtg 2879
|||||
Db 7208 ACCCATTAATAAGCTGTCTCCATTCGCCACTCGGCGCGAGTCTCTACCCCTGCGTGGT 7267
QY 2880 gtacgactgtggccccagcgagcgtgtgaaataaaatcctctgtgtgtgtgatacaaaa 2939
|||||
Db 7268 GTACGACTGTGGCGCCAGCGCGCTTGGAATAAATCCCTTGTGTGTCATCAAAAA 7327
QY 2940 aaaaaa 2945
|||||
Db 7328 AAAAAA 7333

RESULT 14
PEN288592 2000 bp RNA VRL 29-AUG-2000
LOCUS Porcine endogenous retrovirus PERV-B7 env gene for envelope
DEFINITION protein, genomic RNA.
ACCESSION AJ288592
VERSION AJ288592.1 GI:7414453
KEYWORDS env gene; envelope protein.
SOURCE porcine endogenous retrovirus.
ORGANISM Porcine endogenous retrovirus
retroviruses; Retrovirdae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
1 (bases 1 to 2000)
Bosch, S., Arnould, C. and Jestin, A.
Study of full-length porcine endogenous retrovirus genomes with
envelope gene polymorphism in a specific-pathogen-free large white
swine herd
J. Virol. 74 (18), 8575-8581 (2000)
20411432
JOURNAL MEDLINE
REFERENCE 2 (bases 1 to 2000)
AUTHORS Bosch, S., Arnould, C. and Jestin, A.
TITLE Study of full-length porcine endogenous retroviral (PERV) genomes
with envelope gene polymorphism in a large white
specific-pathogen-free swine herd
Unpublished
3 (bases 1 to 2000)
Bosch, S.
Direct Submission
Submitted (28-MAR-2000) Bosch S., Molecular Biology Department,
AFSSA-Zoopole, Zoopole Les Croix B.P.53, Ploufragan 22440, FRANCE
JOURNAL TITLE location/Qualifiers
FEATURES
1..2000
/organism="porcine endogenous retrovirus"
/specific_host="Sus scrofa"
/db_xref="taxon:61673"
/clone="PERV-B7"
/note="French Large White SPF-pig"
1..1899
/gene="env"

gene

CDS
1..1899
/gene="env"
/codon_start=1
/product="envelope protein"
/protein_id="CA86211.1"
/db_xref="GI:7414454"
/translation="MHPILSWRHLPFRGERRRRLIPLSFAFMFLTLTIPOASSK
RLIDSNHRLPLSLTWLIIDPDGTAVNSTGVA PRGTWPELHFLRLINPAVSTP
NLVRSYGFCCPPGTEKREYCGGSGESFCRWSGVTSDGMKMPFLIDRKFSPN
SGPGRYRMKLYKXKSCSPDLVDLKI SFMRKGOENIKDINMSGCIYRYKGGCA
GSTVIIRLRIETGTEPPYAVGPDPVLABOGPPALBEPHNLVPLTISRDPDPSN
GTTGILPTNPNRNSPGVPVKGRKRLFLTQAFKAIKSTDEANSSCWLCLSSPPY
EGMAKEGFNVTKEHRNOCITWGSNRKRLTLEVSQKCTIGAPPSHOPCSTVYBO
ASENOYLPGYNNRMACNTGLPCVSTSVFNQSKDFVMVOIVRYYHYHPEVLDY
DYRNRKPGEPVSLTAVMLGLTAVGCTGTALINGPOOLEGELBHAAMEDIR
ALAESVSLESLTSLSGVVLONRGLDILFLRSGCALAKKRCCTCYVHSGA IRRS
MSKLERLERRRRERRELAQGMFEGNFNSPMTWTLTSLRPLVYVLLLVGCLIN
RFAVARE"
BASE COUNT 532 a 485 c 505 g 477 t 1 others
ORIGIN
Query Match 63.2%; Score 1869.6; DB 14; Length 2000;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1893; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 364 atgcatccacgcttaagctgagcgccacccctccgactcgaggtgagagccgaaagactg 423
Db 1 ATGCATCCACGTTAAGCTGGCGCCACCTTCCGACTCGGCGTGGAGAGCGAANAAGACTG 60
QY 424 agaatcccttaagcttgcctccatgcgctgtgtcttacttaactaactaactccag 483
Db 61 AGAATCCCTTAAGCTTGCGNCTTCCTGCTGCTGCTTCTACTCTPAACAATATATCCACG 120
QY 484 gccagtagtaaacgccttataagacgcctcgaacccccatagaaccttatccctacctg 543
Db 121 GCCAGTAGTAACGGCCTTATAGACAGCTTGAAACCCCACTTATACCTTATCCCTTACCTGG 180
QY 544 ctgattatgacccctgatacggtgtcaactgttaataagcaactcgaggtgtgtcctctga 603
Db 181 CTGATTATTGACCTGATACGGGTGTCACTGTAATAAGCACATCGAGGTGCTTCTCTTGA 240
QY 604 ggcacctggtgctcgaactgcatcttgcctccgatitgataaacccgctgtttaaaagc 663
Db 241 GGCACCTGGTGCGCAAGCACTGCAATTCGCCATGATTAACCCCGCTGTAAAGC 300
QY 664 aaactcccaactagtcgtgtagttagttagttagttagttagttagttagttagttagt 723
Db 301 ACACCTCCCAACCTAGTCCGTAGTATAGGTTCTATTGCTGCCAGGACACAGAAAG 360
QY 724 aaatactgtgggtgttcgaggaatcctctgttagagatgagctgtcaacctcaac 783
Db 361 AATATCTGTGGGGGTTTCGGGAATCTTCTGTAGGAANTGAGCTGCTCACTCCCAAC 420
QY 784 gatgagagactggaatlggcccgaatctctccagagccgggtlaaaaatttccctttgcaat 843
Db 421 GATGAGAGACTGAATATGGCCGAAITCTTCTCCAGACCGGGTAAATTCCTCTTGTCAAT 480
QY 844 tcggccggcgaggaatgacaaatgattgaactatataagataagagctgtctcccatca 903
Db 481 TCCGGCCCGGCGAAGTACGAGCTGTGAACATATATTAAGATTAAGAGCTGCCCAACA 540
QY 904 gacttagattacttaagataagcttcaactgaaagaaagaaacgagaaatattcaaaag 963
Db 541 GACTTAGATTATCTTAAGATTAAGTTTCACTGAAAAAGCAAAACAGAAATATTCTAAAG 600
QY 964 tggataaagtgatagactggggaatagtttttaataatatabgcggggagcgaggttc 1023
Db 601 TGGATAAATAGTATGAGCTGGGGAATAGTTTATTAATATATGCGGGGAGACAGGGTCC 660
QY 1024 acttaaacatcgcccttagatagaagcggggaagaaacccctgtgtgcaatggagacc 1083
Db 661 ACTGTAAACCATTCGCTTAGATAGAGACGGGAGACAGAAACCCCTGTGTGAGTGGAGACC 720

QY 1084 gataaagctgctgagcaagggcccccgccttgagagccagcgaataacttgccggtg 1143
 |||||
 Db 721 GATAAAGTACTGGCTGAACAGGGGCCCCGCCCTGAGGCGACCGCATTAATCTGCCAGTG 780
 QY 1144 ccccaattaaactgcgtgcgcgcctgagatacaacagccgccttagcaacagatacaactgga 1203
 |||||
 Db 781 CCCCCAATTAACTCTGGCTGGGCTGAGACATACAGCCGCCCTTACCAAGGATGCACTGGA 840
 QY 1204 ttgattccctacacacgccttagaactccccaggtgttccctgtttaagacagagagaga 1263
 |||||
 Db 841 TTGATTCCTTACCACATGCTGAGAACTCCCCAGGTGTTCCTGTAAAGACAGGAAGAAGA 900
 QY 1264 ctcttcagttccatccagggagcttccaaagcacaaccacccagccctgatgcact 1323
 |||||
 Db 901 CTTTCAATCTCATCCAGGGAGCTTTTCAAAAGCCATCAAAATCCACCGAGGCCCAAT 960
 QY 1324 tctcttggttgctgtctcctcagggccctcttatattatagggagatgagctaaagaa 1383
 |||||
 Db 961 TCTTCTGTGTGGCTTTGTCTATCTCTCAGGGCCCTCTTATATGAGGGGATGCTAAAGAA 1020
 QY 1384 agaaatccaatgtgacaaagagagacatagaatcaatgatactgaggggtcccgaaataag 1443
 |||||
 Db 1021 GGAATAATTCAATGTGACCAAAAGACCATAGAAATCAATGTACATGGGGGTCCCGAAATAG 1080
 QY 1444 ctacccctcaactgaagttccgggaagggagacatgataagaaagctcccccaaccac 1503
 |||||
 Db 1081 CTTACCTCTCAGGAGATTTCCGGGAGGGAACATGATAGGAAAGCTCCCCCATCCAC 1140
 QY 1504 caaacaccttgtagtactggtgttatagagagagccctcaaaaaactgatttttagta 1563
 |||||
 Db 1141 CAAACACCTTGCTATGACTGCTGCTTATGAGACAGGCTCAGAAATCAGTATTTAGTA 1200
 QY 1564 ccttggtataacaggtgtggtgacatgacatactggttaaacccctgtgttccaccctca 1623
 |||||
 Db 1201 CCTGTTATATAACAGGTGAGTGGCATGCAATAGTGGTTAACCCCGTGTGTTCCACCTCA 1260
 QY 1624 gtcttcaacacatccaaagattgtgtgtacatggttccaaatcgtccccgagtgcttac 1683
 |||||
 Db 1261 GTCTTTTAAACCAATCAAAAGATTTCTGTGATGATGCTCAAAATCGTCCCGAGGTACTAC 1320
 QY 1684 catcctgaggaaggtggtcccttgatgatactgactatcggtatataccagcaaaagagaa 1743
 |||||
 Db 1321 CATCTTGAAGAGAGTGGTCTTGTGATGATATGACTATGCTGATATACCGAACCAGAAAGAGAA 1380
 QY 1744 cccglatcccttaacctagctgttaatgctcgtgattaggaagcggcgttggcgtlaagaa 1803
 |||||
 Db 1381 CCCGTATCCCTTACCTAGCTAGCTGATGCTCGATGCGATTAGGAGCGGCGTGGCGTATGAGAAC 1440
 QY 1804 ggagcaactgcccctgatacagagacacagcagctagagaaagacttggtagctacat 1863
 |||||
 Db 1441 GGGACACCTGGCTGATCAAGGACACACACACTAGAGAAAGACTTGGTGTGACTACAT 1500
 QY 1864 ggcgcctatgacagaagatctccgagcctlaaagagctgtttagcaaccccttagaagagctc 1923
 |||||
 Db 1501 GCGGCCATGACAGAGATCTCCGAGCCTTAGAGAGAGTCTTGTAGCAACCTTAGAAGAGTCC 1560
 QY 1924 ctgactccttctgtcgaagtgtgtctacagaaacggagggagattagatctgctgttctta 1983
 |||||
 Db 1561 CTGACTTCTTWTGCTGAGTGTGCTTACAGAACCGGAGGGAGTATGCTGTGTTCTTA 1620
 QY 1984 agagaaggtggttatgtatgagccttaaaagaagaatgtgtcttcatgtatgatactca 2043
 |||||
 Db 1621 AGAAGAGGTGGTATGTGACGCTTAAAGAAAGAAATGTTGCTTATATGATGATCACTCA 1680
 QY 2044 ggaagcctacagagctccatgaaacaaagcttagaanaaaagcttagaggggctgcaagggaa 2103
 |||||
 Db 1681 GGACCCATCAAGAGACTCATGAGAGAACTTAGAGAAAGTGTAGAGAGGCTTCAAGAGGAA 1740
 QY 2104 agagaaggtgacacaggggtgtgttgaagagatgttcaacaggtctccttggatgacac 2163
 |||||
 Db 1741 AGAGAGGCTGACGAGGGGTGTGTAAGGATGTGTTCAACAGTGTCTCTTGATGACCAAC 1800

QY 2164 ctgcttctgctctgacggggcccccctagtagctgctcctgcttacttaactagttggcct 2223
 |||||
 Db 1801 CTCTTTCCTGCTGTGACGGGACCGCTAGTACCTGCTCTCTGTTACTTACAGTTGGGCT 1860
 QY 2224 tgccttaataaagttgttgccttctgtttagaagacagtagtgcagctcagatcatg 2283
 |||||
 Db 1861 TGCTTATATATATAGTGTGTGCTTGTGCTTTAGAGAAATGATGATGCTCAGATCATG 1920
 QY 2284 gtaacttagggcaac 2296
 |||||
 Db 1921 GTACTTAGACAAC 1933
 RESULT 15
 PEN288588 2000 bp RNA VRL 29-AUG-2000
 LOCUS Porcine endogenous retrovirus PERV-AB7 env gene for envelope
 DEFINITION protein, genomic RNA.
 ACCESSION AJ288588.1 GI:7414445
 VERSION env gene; envelope protein.
 KEYWORDS porcine endogenous retrovirus.
 SOURCE porcine endogenous retrovirus
 ORGANISM Viruses; Retrovirdae; Retroviridae; Mammalian type C
 retroviruses; 1 (bases 1 to 2000)
 REFERENCE 1 (bases 1 to 2000)
 AUTHORS Bosch,S., Arnauld,C. and Jestin,A.
 TITLE Study of full-length porcine endogenous retrovirus genomes with
 envelope gene polymorphism in a specific-pathogen-free large white
 swine herd
 JOURNAL J. Virol. 74 (18), 8575-8581 (2000)
 MEDLINE 20411432
 REFERENCE 2 (bases 1 to 2000)
 AUTHORS Bosch,S.
 TITLE Direct Submission
 JOURNAL Submitted (28-MAR-2000) Bosch S., Molecular Biology Department,
 AFSSA-Zoopole, Zoopole Les Croix B.P.53, Ploufragan 22440, FRANCE
 FEATURES
 source
 1..2000
 /organism="porcine endogenous retrovirus"
 /specific_host="Sus scrofa"
 /db_xref="taxon:61673"
 /clone="PERV-AB7"
 /note="French Large White SPF-pig"
 1..606
 /gene="env"
 1..606
 /note="truncated, not functional"
 /gene="env"
 /codon_start=1
 /product="envelope protein"
 /protein_id="CAB86207.1"
 /db_xref="GI:7414446"
 /translation="MHPTLSMRHLPTRGSEPRRLRPLSPASTAFLTLTTPQASSK
 RLIDSSNHPRLISITWLTIDPDGVTVNPSGVAPRGTWMPDELHFCRLINPAKSPR
 PNLVRSYFTCCPGETEKRTKGGGSEFGRMSGVTSNDGDKMPLISLDQVRFSPV
 SGPGRKYMKLYKDKSCSPSDLDLKISFTYEKGRKLFKSG"
 BASE COUNT 532 a 496 c 499 g 473 t
 ORIGIN
 Query Match 63.2%; Score 1868.4; DB 14; Length 2000;
 Best Local Similarity 98.6%; Pred. No. 0;
 Matches 1906; Conservative 0; Mismatches 26; Indels 2; Gaps 2;
 QY 364 atgcctccaggttaagctgagggcactcccgactggggttgagagcgaagaaagactg 423
 |||||
 Db 1 ATGCATCCACGTTTAAGCTGGGCGCACTCCCGACTCGGGGTGAGAGCCCAAAAGGCTG 60
 QY 424 agaattcccttaagctcgctccatcgctgtgttcttacttaacttaactactccag 483
 |||||
 Db 61 ACAAATCCCTTAAGCTTGCCCTGCAATGCCCTGCTTCTTACTCTAACAATAACTGCCAG 120

Mon Feb 25 07:44:31 2002

us-09-171-553b-9.feb22std.rge

```

|||||
401 GAGGCTGCGTCCACCTCCAAACGATGGAGACTGGAAATGGCCGATCTCTC 450
151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrLy 167
451 CAGGACGGGGTAAATTTCTCTTTGTTCATTCCTCCGGCCGGGACAGTACA 500
167 SmetMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAspT 184
501 AATGATGAAGCTATATAAAGATAGAAGCTGCCCATCAGACTTATGATT 550
184 yLeuLysLLeSerPheThrGluLysGlyLysGlnGluAsnTLeGlnLys 200
551 AACTAAGATTAAGTTTCACTGAAAAAGAAAAAGAAAAATTTCAAAAG 600
201 TTPLeasnglyMetSerTrpGlyLLeValPheTyrLysTyrGlyGlyG 217
601 TGGATAAATAGTATGAGCTGGGGAATAGTTTATTAATATAGCGGGGG 650
217 yAlaGlySerThrLeuThrLeaArgLLeGluThrGlyThrGluP 234
651 AGCAGGGTCCACTTACCATTCGGCTTAGATAGACGGGAGACAGAAC 700
234 roProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProPro 250
701 CCCCTGTGGCAATGGGACCCGATTAAGTACGTGCTGAACAGGGCCCCG 750
251 AlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuAr 267
751 GCCCTGGAGCCACCCCATACCTTGGCGGTGCCCAATTAACCTCGCTCG 800
267 gProAspLLeThrGlnProProSerAsnSerThrGlyLeuLeuProT 284
801 GCCTACATTAACACAGCCGCTAGACAGTACCATCTGATTGATTCTCTA 850
284 hrAsnThrProArgAsnSerProGlyValProValLysThrGlyGlnArg 300
851 CCAACACCCCTAGAAACTCCAGGTGTTCTGTTAAGACAGGACAGAGA 900
301 LeuPheSerLeuLLeGlnGlyAlaPheGlnAlaLeuAsnSerThrAspPr 317
901 CTCTTCACTTCATCCAGGGAGCTTCCAGCCATCAACTCCACCGACCC 950
317 oAspAlaThrSerSerCysTrpLeuGlyLeuSerSerGlyProProTyr 334
951 TGAAGCCCACTTCTCTTGTGCTTGTCTTGTCTTGTCTTGTCTTGTCTT 1000
334 yTGluglyMetAlaLysGlnArgLysPheAsnValThrLysGlnHisArg 350
1001 ATGAGGGGATGGCTAAAGAAAGAAATTCATGTGACCAAGACATAGA 1050
351 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSe 367
1051 AATCAATGTATCATGGGGGTCCGAAATTAAGCTTACCTCATGAAAGTTTC 1100
367 rGlyLysGlyThrCysLLeGlyLysAlaProProSerHisGlnHisLeuc 384
1101 CGGGAAGGGGACATGCATAGAAAAAGTCCCCCATCCACCAACACCTTTT 1150
384 yTyrSerThrValValTyrGlnGlnAlaSerGlnAsnGlnTyrLeuVal 400
1151 GCTATAGTACTGTGTATGAGAGAGCCCTCAGAAATCGATTATTAAGTA 1200
401 ProGlyLysAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysVa 417
1201 CCTGGTATTAACAGGTGGTGGCATGCAATACGTGGGTAAACCCCTGTGT 1250
417 lSerThrSerValPheAsnGlnSerLysAspLeuGlyValMetValGlnI 434
1251 TTCACCTCAGTCTTCAACCAATCCAAAGATTGTGTGTCATGTGTCCAAA 1300
434 lValAlProArgValTyrTyrHisProGlnGluValValLeuAspGluTyr 450
|||||
|||||
1301 TCGTCCCCCGAGTGTACTACCATCTGAGAAAGTGTCTTGTATGAATAT 1350
451 AspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAl 467
1351 GACTATCGGTATTAACCGACAAAGAAAGAACCCGTATCCCTTACCCTAGC 1400
467 aValMetLeuGlyLeuGlyThrAlaValAlGlyValGlyThrGlyThrAla 484
1401 TGTATGTGTGGATTAGGAGAGGGCGCTTGGCGTATGAGAACAGGAGACTG 1450
484 lAlaLeuThrGlyProGlnGlnLeuGlnLysGlyLeuGlyLeuHis 500
1451 CCCGATCACAAGGACACAGCAGCTAGAGAAAGACTTGGTGAACAT 1500
501 AlaAlaMetThrGluAspLeuArgAlaLeuLysGlnSerValSerSne 517
1501 GCGGCCATGACAGAGATTCCTCCAGCTTAAAGAGATCTGTAGCAACCT 1550
517 uGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 534
1551 AGAAGAGTCCCTGACTCTTGTCTGAAAGTGTTCACAGAACCGAGGG 1600
534 lLysLeuAspLeuLeuPheLeuArgGlyGlyLeuGlyLysAlaLeuLys 550
1601 GATTAGATCTGTCTTCTTAAGAAAGTGGCTTATGTGACGCTTAAAA 1650
551 GluGluCysCysPheTyrValAspHisSerGlyAlaLeuArgSperme 567
1651 GAAGAAATGTTGCTTCTATAGATCACTCAGAGGCCATCAGAGACTCCAT 1700
567 lAsnLysLeuArgLysLysLeuGlnArgArgArgGluArgGluAla 584
1701 GAACAAAGCTTAGAAAAAGTTAGAGAGCGCTCAAGAGGAAAGAGAGCTG 1750
584 spGlnGlyTrpPheGlyGlyTyrTrpPheAsnArgSerProTrpMetThr 600
1751 ACCAGGGGTGGTTTGAAGATGGTTCAACAGGCTCTCTTGGATGACACCC 1800
601 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeu 617
1801 CTGCTTCTGCTCTGACGGGGCCCTAGTAGTCTCTCTCTTACTTAC 1850
617 rValGlyProCysLeuLLeAsnArgPheValAlaPheValArgGluArgV 634
1851 AGTTGGCCCTGCTTAAATTAATAGGTTGTCTCTTGTATAGAAAGAG 1900
634 alSerAlaValGlnLeuMetValLeuArgGlnGlnTyrGlnGlyLeuLeu 650
1901 TGAATCAGTCCAGATCATGTACTTATGAGCAACAGTACCAAGGCTTCTG 1950
651 SerGlnGlyGluThrAspLeu 657
1951 AGCCAAAGGAGAAACTGACCTC 1971
seq_name: gb_pat:AX002804
seq_documentation_block:
LOCUS AX002804 3482 bp DNA PAT 21-AUG-2000
DEFINITION Sequence 3 from Patent WO9853104.
ACCESSION AX002804
VERSION AX002804.1 GI:9885132
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 3482)
AUTHORS
Stoye,J.P. and Weiss,R.A.
TITLE
Detection of retroviral subtypes based upon envelope specific
sequences
JOURNAL
Patent: WO 9853104-A 3 26-NOV-1998;
MEDICAL RES COUNCIL (GB); STOYE JONATHAN PAUL (GB)
FEATURES
Location/Qualifiers

```

source 1..3482
 /organism="porcine endogenous retrovirus"
 /db_xref="taxon:61673"
 BASE COUNT 927 a 854 c 867 g 834 t
 ORIGIN

alignment_scores:
 Quality: 3484.00 Length: 657
 Ratio: 5.319 Gaps: 0
 Percent Similarity: 99.696 Percent Identity: 98.478

alignment_block:

US-09-171-553b-10 x AX002804

Align seg 1/1 to: AX002804 from: 1 to: 3482

1 MethiSProThrLeuSerTrpArgHisLeuProThrArgGlyGlyLeuPr 17
 911 ATGCAATCCACCTTAAGCTGGCGCCACTCCCGACTCGGGGTGGAGAGCC 960
 17 OlySarGluLeuArgLeuProLeuSerPheAlaSerIleAlaTrpPheLeu 34
 961 GAAAGACCTGAGAAATCCCTTAAGCTGCGCTCCATCGCTGGTCCCTTA 1010
 34 hLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50
 1011 CTCTAACATAATACCTCCAGGCACTAGTAACGCTTATAGACAGCTCG 1060
 51 AsnProHisArgProLeuSerProThrTrpLeuIleIleAspProAspTh 67
 1061 AACCCCATAGACCTTATCCCTTACCTGCGGTGATTTGACCCGTGATAC 1110
 67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84
 1111 GGGTGCACTGTAATAGCAGCTCGAGGTGCTGCTAGAGGACACCTGCT 1160
 84 rProGlnLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
 1161 GGCCTGAACGATTCCTGCTCCGATGATTACCCGCTGTATAAACGC 1210
 101 ThrProAsnLeuValArgSerTrpGlyPheTrpCysCysProGlyTh 117
 1211 ACACCTCCCAACCTAGTCGCTAGTATGGGTCTATGCTGCCAGGCAC 1260
 117 rGlyLysGlyLysTrpCysGlyGlySerGlyLysLeuSerPheCysArgArg 134
 1261 ACAGAAAGAGAAATACTGCTGGGGGTCTGGGGAATCCTTCTAGAGAT 1310
 134 rPserCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150
 1311 GGAAGCTGCTCACTCCAAAGATGAGAGTGAATGGCCGATCTCTCTC 1360
 151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTrp 167
 1361 CAGGACCGGGTAAATTCCTTGTGCAATCCGCGCCGCGCAAGTACAA 1410
 167 sMetMetLysLeuTrpLysAspLysSerCysSerProSerAspLeuAsp 184
 1411 AGTGATGAACATATATMAAGATAGAGCTGCTCCCATCAGACTTAGATT 1460
 184 rLeuLysLysLeuSerPheTrpGlyLysGlyLysGlnGlnLysnIleGlnLys 200
 1461 ATCTAAAGATAGATTCTCACTGAAAAAGAAAAACAGAAAAATATTCAAA 1510
 201 TrpIleAsnGlyMetSerTrpGlyIleValPheTrpLysTrpGlyGly 217
 1511 TCGATTAATGGTATGAGCTGGGGAATAGTTTTTATATAATATGCGGGGG 1560
 217 yAlaGlySerThrLeuThrIleArgLeuArgIleGluTrpGlyThrGluP 234
 1561 AGCAGGGTCCACTTAACTTGGCTTAGATAGAGAGCGGGACAGAAC 1610

234 rObroValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProPro 250
 1611 CCCCTGGGAGAGTGGGAGCCCGATAAAGTACTGCTGCAACAGGGGCCCCG 1660
 251 AlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeu 267
 1661 GCCCTGGAGCCACCGCATTAACCTTCGGGTCCCAATTAACCTCGCTGGC 1710
 267 gProAspIleThrGlnProProSerAsnSerThrThrGlyLeuIlePro 284
 1711 GCCTGACATACACAGCCGCTTACCAAGCTTACCTGAGGCTTCCTTAT 1760
 284 hAsnThrProArgAsnSerProGlyValProValLysThrGlyGlnArg 300
 1761 CCAACAGCCTAGAAACTCCCAAGGTGCTCTTATAGACAGAGAGAGA 1810
 301 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAsp 317
 1811 CTCTTCAGTCTCATTCCAGGAGCTTTCAGCATCAACATCCACCGACCC 1860
 317 cAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTrp 334
 1861 TGATGCCACTTCTTCTTGTGGCTTTGTCTATCTCAGGGCTTCCTATT 1910
 334 yrgLuglyMetAlaLysGluArgLysPheAsnValThrLysGluHisArg 350
 1911 ATGAGGGGATGGCTAAGAAAGAAATTCATGACGACCAAGAGCATAGA 1960
 351 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluVal 367
 1961 AATCAATGATACATGGGGGCTCCGAAATAGCTTACCTCACTGAAAGTTTC 2010
 367 rGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeu 384
 2011 CGGGAAGGAGCATATGCAATAGAAAGCTCCGCCATCCACACACACTT 2060
 384 ysrYrSerThrValValTrpGluGlnAlaSerGluAsnGlnTrpLeuVal 400
 2061 GCTATAGACTGTGTTATAGACAGGCTCAGAAATCAGTATTATAGTA 2110
 401 ProGlyLysAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysVa 417
 2111 CCTGTATATAACAGGTGGTGGCATGCAATACGCGGTAAACCCCTGTGT 2160
 417 lSerThrSerValPheAsnGlnSerLysAspLeuValMetValGln 434
 2161 TTCCACTCAGTCTTCAACCAATCCAAAGATTCTGTGTCATGCTCCAAA 2210
 434 lValProArgValTrpTrpHisProGluGluValValLeuAspGluTrp 450
 2211 TCGTCCCGCCGAGTGTACTACCATCTGAGAAAGTGTCTTGATGATAT 2260
 451 AspTrpArgGlyAsnArgProLysArgGluProValSerLeuThrLeuAl 467
 2261 GACTATCGGTATTAACGACCAAAAGAGAACCCCTATCCCTTACCTAGC 2310
 467 aValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrAla 484
 2311 TGATATGCTGGATTAGGAGCGCGCTTGGCTGAGAAACAGGACAGCTG 2360
 484 lAlaLeuIleThrGlyProGlnGlnLeuGlyLysGlyLeuGlyLys 500
 2361 CCGTATATCAGAGACACAGCAGCTAGAGAAAGAGACTTGGTGAAGCTCAT 2410
 501 AlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnLe 517
 2411 GCGGCATGACAGAGAGATCTCCGAGCTTAAAGAGAGTGTGTACCACT 2460
 517 uGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 534
 2461 AGAAGAGTCCCTGACTTCTTGTCTGAAGTGTCTTACAGAAACGAGGAG 2510
 534 lYleuAspLeuLeuPheLeuArgGluGlyLysGlyLysAlaAlaLeuLys 550

2511 GATTATGATCGCTGTTTCTAAGAGAGGTGGGTATGTCAGCCTTAAAA 2560
|||||
551 GluGluCysCysPheTyrValAspHisSerGlyAlaLeuArgAspSerMet 567
|||||
2561 GAAGAATGTTGCTTCATCTATGATCATGCACTCAGGAGGCATCAGACACTCAT 2610
|||||
567 TAAuLyLLeuArgLyLysLeuGluAbArgArgArgArgGluArgGluAla 584
|||||
2611 GACCAACGCTTAGAGAAAGGTTTGAAGGGCGTGTGAAGGGGAAAGAGGCGTG 2660
|||||
584 spGlnGlyTyrPheGlnGlyTyrPheAsnArgSerProTribethrThr 600
|||||
2661 ACCAGGGGTGCTTTGAAGGATGCTTCACAGGTCCTTGATGACCCACC 2710
|||||
601 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeuTh 617
|||||
2711 CTGCTTTTCGCTCTGACGGGAGCCCAAGTAGCTCGCTTCCTTACTTAC 2760
|||||
617 rValGlyProCysLeuLeuAsnArgPheValAlaPheValArgGluArgV 634
|||||
2761 AGTGTGGCGCTTGCTTAATTAATAGGTTTGTCCTTTGTTAGAGACAG 2810
|||||
634 AlSerAlaValAlaGlnIleMetValLeuArgGlnGlnIleTyrGlnGlyLeu 650
|||||
2811 TGAGTGCATCCAGATCATGACTTAAAGCAACAGTACCAAGGCGCTTCTG 2860
|||||
651 SerGlnGlyLysThrAspLeu 657
|||||
2861 AGCCAAAGAGAACTGACCTC 2861

seq_name: gb_v1:PERENV2

	seq_documentation_block:		VRL	14-OCT-1997
LOCUS	PERENV2	3482 bp	RNA	
DEFINITION	Porcine endogenous retrovirus env gene,	3482 bp.		
ACCESSION	F12239			
VERSION	y12239.1	GI:2576324		
KEYWORDS	env gene.			
SOURCE	porcine endogenous retrovirus.			
ORGANISM	porcine endogenous retrovirus.			

REFERENCE	1 (pages 1 to 3482)
AUTHORS	Le Tissier, P., Stoye, J. P., Takeuchi, Y., Patience, C. and Weiss, R. A.
TITLE	Two sets of human-tropic pig retrovirus
JOURNAL	Nature 389 (6652), 681-682 (1997)
MEDLINE	97478526
REFERENCE	2 (pages 1 to 3482)

TITLE	Direct Submission
JOURNAL	Submitted (02-APR-1997) J.P. Stoye, National Institute for Medical Research, Virology, The Ridgeway Mill Hill, London, NW7 1AA, UK
REMARK	revised by author 16-SEP-97
FEATURES	Location/Qualifiers
Source	1. .3482

gene
CDS

```
/codon_start=1
```

/db_xref="GI:2576325"

```
/translation="MHP.TLSWRHLP
```

PNLVRSYGFYLCPPGTEKEKYCGGSGE
SGPGKYKMKI.YRDKSCSPSDI.DYI.K

..GJLI INKRIEIGIEPVAAGPDKVL

```

GTTGCTLTPTNTPRNSPGVPVYTCGRLSELTIGAGAOAINSTDPPDATSSCWLCLSSGPYY
EGNAKEKEFNVTKEHNRQCTWGSRNKTLTLEVSCKGCTICIKAPSPHQHLCSTVVEQ
ASENQYLVPGYNRMWMACTNGTLTPCVSTSVNSKDFCVMWOIVPRVYVHPEEVLDEY
DYRNRNRKREPEVSLTLLAVMLGCTAVGVGTGALLITGPOLKEKGLGELHAAMTEDLR
ALBEYSRNLSEESTLTSLEVVLQNRGRGIDILFELREGGICALKECCGYPVHSGAIRDS
MSKIRLELRRRRRRREADOGCFEGMGFSRPMWTTLLSALGPLVLLLLLTIVGPCLIN
RFAFAVERERSAVOIMLRQOYGLLSGSETDL"

BASE COUNT      927 a      854 c      867 g      834 t
ORIGIN

alignment_scores:
    Quality: 3484.00      Length: 657
    Ratio: 5.319          Gaps: 0
    Percent Similarity: 99.696      Percent Identity: 98.478

alignment_block:
US-09-171-553B-10 x PERENV2      ..
Align seg 1/1 to: PERENV2 from: 1 to: 3482

```

Align seg 1/1 to: PERENV2 from: 1 to: 3482

1 MethisProThrLeuSerTTPargHisLeuProThrrargGlyGlyGlyIupr 17

311 ATGCAATCCACGTTAAGCTGGCCGACCTCCGACTCGGGGTGGAGACC 96

[illegible]

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

111

51 ASNPROHISARRDPROLEUSERRDROTHRTYRLGUTLQTLQANDDVALSSEH 67

661 AACCCTATAGACCTTTATCCCTTACCTGGCTGATTATTGACCTGATAC 11

67 rGlyValThrValAsnSerThrArgIvValAlaProArgIvThrTrpT 84

111 GGGTGTCACTGTAATAGCACTCGAGGTGTTGCTCTCCTAGAGGCACCTGGT 111

84 rpprogluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100

161 GGCCTGAAC TGCA TTTCTG CCTCCGATTG ATTAACCCCGCTGTTAAAGC 12

101 **IN**ProRoASHLeuValArgSerIyrGlyPheIyrCysCysProGlyTh 11

[illegible][illegible]

111

51 G]naSnArvA]IwSbheSerDhva]aenSorC]uBroC]uIwTut.. 16

361 CAGGACCGGTTAAATCTCTCTTGTCAATTCGGCCCGGGCACTACAA 14

667 smetmetIysleuTYrIlySAspLySSerCySSerProSerAspLeuAspT 18

11 AGTGAATGAACCTATATAAGATAAGAGCTGCTCCCATCAGACTTAGATT 14

84 yrLeuLysIleSerPheThrGluLysGlyLysGlnGluAsnIleGlnLys 20

61 ATCTAAGATAAGTTTCACTGAAAGAAGCAAAACAGGAAATATTCAAAG 15.

TrpLeasnGLymetSerTrpGlyIleValPheTyrLysTyrcylcylcyl 21

11 16GATAAATGGTATGAGCTGGGAAATAGTTTATTATAAATATGGCCGGGG 156

```

2461 AGAAGAGTCCCTGACTCTTGTGCTGAATGCTGTACAGAAACCGAGGG 2510
|||||
534 lyeuaspLeuApuheLeuAargLugLyLyLeuCyAlaAlaLeuLys 550
|||||
2511 GATTAGATCTGCTGCTTCTTAAGAGAAAGGTGGGTATGTGCAGCCCTAAAA 2560
|||||
551 GluGluCyScsPheLyValAspHisSerGlyAlaIleAArgpSerMe 567
|||||
2561 GAGAGATGTGCTTCTTATGTAGATCACTAGAGAGCCATCAGAGACTCAT 2610
|||||
567 LasnLysLeuAryLysLysLeuGluAArgAArgAArgGluAargGuaAla 584
|||||
2611 GAGCAAGCTTAGAGAAAGCTTAAAGAGCCGCTGAAGAGGAAAGAGAGCGTG 2660
|||||
584 spGInGlyTTPPhGluGlyTTPPhAsnArgSerProTrrpMethTrThr 600
|||||
2661 ACCAGGGGTGGTTGAAGAGATGGTTAACAGGGTCTCTGGATGACACGC 2710
|||||
601 LeuLeuSerAlaLeuThrArgGlyProLeuValAlaLeuLeuLeuLeuThr 617
|||||
2711 CTGCTTCTGCTGTGACGGGAGCCCTAGTACGCTGCTGCTGTACTTAC 2760
|||||
617 rValGlyProCysLeuIleAsnAryPheValAlaPheValAargGluAry 634
|||||
2761 AGTTGGGCTTGGCTTAATTAATAGGTTTGTGGCTTTGTTAAGAGACGAG 2810
|||||
634 aLseAlaValAlaGlnIleMetValLeuAArgGlnIlyrGInGlyLeuLeu 650
|||||
2811 TGAGTGCAGCTCCAGATCATGTACTTAAAGCAACAGTACCAAGGCTTGTG 2860
|||||
651 SerGInGlyGluThrAspLeu 657
|||||
2861 AGCCAAAGAGAAACTGACCTC 2881
|||||
seq_name: gb_v1: PEN293657

seq_documentation_block: 8763 bp DNA VRL 30-MAY-2001
LOCUS PEN293657
DEFINITION Porcine endogenous type C retrovirus proviral gag gene, pol gene
and env gene, class B, clone 213.
ACCESSION AJ293657
VERSION AJ293657
KEYWORDS env gene; envelope; gag gene; group specific antigen; pol gene;
polymerase; polypoteIn..
SOURCE porcine endogenous type C retrovirus.
ORGANISM porcine endogenous type C retrovirus
VIRUSES; Retrovird viruses; Retroviridae; Gammaretrovirus.
REFERENCE 1 (bases 1 to 8763)
AUTHORS Krach,U., Fischer,N., Czanderu,F. and Tonjes,R.R.
TITLE Comparison of replication-competent molecular clones of porcine
endogenous retrovirus class a and class b derived from pig and
human cells
JOURNAL J. Virol. 75 (12), 5465-5472 (2001)
MEDLINE 21256017
REFERENCE 2 (bases 1 to 8763)
AUTHORS Toenjes,R.R.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-2000) Toenjes R.R., Medical Biotechnology,
Paul-Ehrlich-Institut, Paul-Ehrlich-Str. 51-95, Langen, 63225,
GERMANY

FEATURES
source location/Qualifiers
1..8763
/organism="porcine endogenous type C retrovirus"
/specific_host="Sus scrofa"
/db_xref="taxon:70540"
/clone="213"
/cell_line="PK15"
/note="class B"
1..630
LTR
misc_feature 467
/note="Cap site"

```

```

primer_bind 633..650
gene 1077..2651
/ gene="gag"
CDS 1077..2651
/ gene="gag"
/ codon_start=1
/ product="group specific antigen"
/ protein_id="CAC39625.1"
/ db_xref="GI:14275842"
/ translation="MGQVTPPLSLTLDHWTENKRAHNLSVYKKKQWQTCASEMP
TEVDGWPSEGFENSEILLAVKAILFQTPGSHPOEYVLTMDLAEPPWPKM
KPRKPRILALIEKNKSHAEKVPSPGIYELIEPTWPEPPVPPADGAVRG
PSAPGAPVVEGPAAGTSSRGATPERTDILILPLTGYPPMGGLQIQWPPSS
ADLYNKTNNPPESEDPORLGLVESLMEHQTPWDCOOLITLTTEREILLEA
KKNVPGADGRPTOLQNEIDMGFLTRPGMDYNTEGESLKITRQALVALKASRRP
TNLAKREVWQGPNEPPSVFLERLMEAFRRFPDPTSEOKASVALAFIQSALDIR
KTLRLGLOEALRDLREAEKYYRRETEEEKQGEKREERERDRRQENKIL
KTLAAVEGKSSREDFPKIRSGPQSGNLNRPFLDKQCAVCYCKEKGMAKRNCPK
KGNKGFVALLEEDKD"
gene 2652..6239
/ gene="pol"
/ gene="pol"
/ gene="pol"
/ note="complete pol ORF, gag-pol is expressed as a
polyprotein"
/ citation=[1]
/ codon_start=1
/ product="polymerase"
/ protein_id="CAC39626.1"
/ db_xref="GI:14275843"
/ translation="GRRGSDPLPEPRVTLKVEGQVPEFLVDGAEHSYLLQPLGRIKE
KSWMGATGQROYPTTRRTVDLANGVTHSELVIECPVLLGDLTKKGAQISF
EGCRPEYSVNNKPTIVTLQLDDEYRLYSQVPPDDISMLEQFOAEAGMGLA
KQVPPVQTLKASATPSVQRYLRSAREGQWIPVORLQOGLIIVYOSPMNTPLP
VRPGNDYRPFVODLREVNKRVODIHPVNPNNLSALPEPENNVTYLDLDAFEC
RLHPTSQPLFAFEPDRDQPGTGQTLTWTRLPQGFKNKSPFPEALRIHOLANRIOHQ
VTLQYVDLILAGATKODCLBGTALLLEISDGRASAKKAQICRRVYTLGSLK
GGORMLTEARKKTYVOIIPATTAQKQVRETLTAGFRMLIPGATLAALVPLTEKRG
ERSWAPBHQKAPATIKALISAPALAPVTKPFLYDERGVAGVLTQILGMRK
PVATLSKRLDPAVSGMPCLKAAVAAILVKADKLTLCQNTVLAHALNEMIVQRP
DRWMTNARMTYOSLLTERVTEFAPPALNPPATLLPEETDEVTDCOHLLEETGVR
KDLTDIPLGEVLTWFTDGSYVEGKMAAGVADGTITVASSIPRESTGAKRME
ALTOALRLAEKSSINITYDSRYAFATAHVGALIVYORGLITSAGREIKKEISILE
ALHLPKRALIHCPOHOKAKDLISRNOMADYAKAOAAVANVLLPIETPKAPEPRQ
YTIEMQELIKIDQSETEGTCTITSYKEILLPHKEGLYVQIHLRLHGLKHLQOL
VRTSPYHLRLPGVADSVVHKVPCQLVANPSRIIPGKRLGSHDGAHWEVDTEVK
PAKYGNKYLLLEVDFESGWVEAYPTKETSIVAKKILEEIPPRGIPKVISDNGPA
FAVOVSQGLAKILGIDMKLCAVRRQSGOVBRMNTIKETLTKLTETGIDMTAIL
PVLPRVNTPGOGLTPEYELLYGGRPLVETASVSHADVLLSOPLSFKRLAEVRO
RAMKOLREAYSGEDQVPHRYQVDSYVRRHHHGLETRWKGYVLLTLTPYAVKY
EGISTWIIASHVKLAPPDGSMRAEKTENPKLRLHRLVPSNNNSPQO"
gene 6112..8085
/ gene="env"
/ gene="env"
/ gene="env"
/ gene="env"
/ codon_start=1
/ product="envelope"
/ protein_id="CAC39627.1"
/ db_xref="GI:14275844"
/ translation="MHPITLSMRHLPTRGGEPRRLRIPSPASIAWFLTLTTPOASSK
RLIDSNPRLSLTWIILDPDGTVNSGTAVARGWMPHNLCLILINPAKSTP
PMLVSYGVCOPPEKEKCGSGESCRKRSCTSDGDMKMFISIDVRKFSFN
SGPGKWKWKLLKDKSCFSIDLIVLKISFTEKQENLOKMINGMSWIVKYGGA
GSTTLIRLKEETEPVAVAGDGLAEOGPALPEPNLIPVPLTSLRPDITPSYX
STTGILPNTSPNSPGVPGKGLFSLQGAFOALINSTDDATSCULCLSSGPPY
EGMAKEGFENVTKERNOCITWGSRNKILFTEVSGGCTIGAPSHOHLCSTVYED
ASENOYLGVGNRMWACNTGLPCVSTSVFNOQFCOMQIVPVPVYHPREPVDE
DYRYNRPRREPSTLAVMLIGLTAVGCTAALITPQDEKGLGHLHAMTEDLR
ALEEVSUNLESITSLSEVYLNRRGLDLTLRBSGLCAALKERCCFVDSGAIIRS
MSKLERLERRRERREADQGWFEQGFNSPMTTLLSALTGPLVLLLLLVGCLLIN
REVAVEREVSVAQIMVLRQOYOGILSOGETDL"
LTR 8133..8763
polyA_signal 8655..8660

```

```

BASE COUNT 2301 a 2200 c 2287 g 1974 t 1 others
ORIGIN
alignment_scores:
  Quality: 3472.00
  Ratio: 5.317
  Percent Similarity: 99.391
  Percent Identity: 98.326
alignment_block:
US-09-171-553B-10 x PEN293657
Align seg 1/1 to: PEN293657 from: 1 to: 8763
1 MethisProthrusSerTriPARHISleuProthrusArglyGlyIupr 17
|||||
6112 ATGATCCACAGTTAAAGCTGGCGCACCTCCGACTGGGCTGAGAGGCC 6161
17 OlyargLeuArgIleProLeuSerPheAlaSerIleAlaTriPheLeu 34
|||||
6162 GAAAGACTGAGATCCCTTAAGCTTCGCTCATCGCTGCTGCTTCTTA 6211
34 hrieuthrIlethrProGlnAlaSerIlyArgLeuIleasPserSer 50
|||||
6212 CTCTAACAAATAACCCAGGAGAGTAAAGCCCTTATAGACACTCG 6261
51 AsnProHisArgProLeuSerProthTriPheIleleasProHis 67
|||||
6262 AACCCCATGAGACTTATATCCCTTACCTGCTGCTATATATGACCTATAC 6311
67 rGlyValThrValAsnSerThrArgIlyValAlaProArgIlyThrTriP 84
|||||
6312 GGGGTGACATGTAATAGACCTGAGGTGTGCTCTAGAGGACACTGGT 6361
84 rPrProGlyLeuHisPheCysLeuArgLeuIleasnProAlaValIysSer 100
|||||
6362 GGCTTGACACGATTTCTGCTCGATGATTAACCCGCTGTATAAACG 6411
6362 TheProGlnLeuValArgSerIlyGlyPheThrCysCysProGly 117
|||||
6412 ACACCTCCCACTAGTCCGTACTTATGGCTTATATGCTGCCAGGAC 6461
117 rGlyValGlyIlyThrCysGlyIlySerGlyIlySerPheCysArgArg 134
|||||
6462 AGAGAAAGAAATACTGTTGGGGGCTTCTGGGGAATCTTCTAGAGAG 6511
134 rPserCysValThrSerAsnAspGlyAspTriPlyStrProIleSerLeu 150
|||||
6512 GAGCTGCGGTACCTCAACGATGAGACTGGAAATGCGGATCTCTCTC 6561
151 GlnAspArgValIlyPheSerPheValAsnSerGlyProGlyIlyStrIly 167
|||||
6562 CAGGACCGGTAATAATCTCTTGTGCAATTCGCGCGCGGCAAGTACAA 6611
167 sMetIlyIlySerIlyAspIlySerCysSerProSerAspLeuAsp 184
|||||
6612 AGTGATGAATAATATAAGATAGAGCTGCTCCCATCAGACTTAAAT 6661
184 yIleuIlyIleSerPheThrGlyIlyGlyGlnGlnAsnIleGlnIly 200
|||||
6662 ATCTAAAGATAAGTTTCACTGAAAAAGAAACAGAAAAATATCTCAAA 6711
201 TrpIleAsnGlyMetSerIlyGlyIlyValIlePheIlyIlyStrIlyGly 217
|||||
6712 TGGAATAAGATATGATGCTGGGGAATAGTTTTTTTAAATATAGGGGG 6761
217 yAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGlu 234
|||||
6762 AACAGGCTCACTTAACCATTCGCTTAGATAGAGAGGAGAGAGAAC 6811
234 rOPrOValAlaMetGlyProAspIlyValLeuAlaGluGlnGlyProPro 250
|||||
6812 CCCCTGCGCAGTGGGACCGATTAAGTACGTGAGAACAGGCGGCCCG 6861

```


251 AlaLeuGlProHisAsnLeuProValProGlnLeuThrSerLeuR 267
 |||||
 6862 GCCCTGGAGCCACCCGATTAAGTGGCGGCGCCCAATTAACTCCCTGGC 6911
 267 gProAspIleThrGlnProProSerAsnSerThrThrGlyLeuIleProT 284
 |||||
 6912 GCGTGACATACACAGCCGCTTAGCANCAGTACACTGGATGATGATTCCTA 6961
 284 hrAsnThrProArgAsnSerProGlyValProValIleThrGlyGlnArg 300
 |||||
 6962 CCACAGCGCTAGAAACTCCCGAGTGTCTCGTTAAGACAGACAGACAGA 7011
 301 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspR 317
 |||||
 7012 CTCTTCAGTCTCATCCAGGAGCTTTCACAGCATCACTCAACGCGACCC 7061
 317 oAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTyrT 334
 |||||
 7062 TGATGCCACTTCTTCTTGTGGCTTGTCTATCTCAGGGCCCTTATT 7111
 334 yrlGlnLysMetAlaLysGlnArgLysPheAsnValThrLysGlnHisArg 350
 |||||
 7112 ATGAGGGGATGGCTAAAGAAAGAAATCAATGTGACCAAGAGCATAGA 7161
 351 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGlnValSe 367
 |||||
 7162 AATCAATGATACATGGGGGCTCCGAAAVAAAGCTTACCTCATCGAAGTTTC 7211
 367 rGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeu 384
 |||||
 7212 CGGGAAAGGGACATGCATAGAAACTCCCGCATCCACCAACACTTT 7261
 384 ysfYrSerThrValValTyrGlnGlnAlaSerGlnAsnGlnTyrLeuVal 400
 |||||
 7262 GCATAGTACTGTGTTATAGACAGAGCCCTCAAGAAATCAGTATTAGTA 7311
 401 ProGlyLysAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysVa 417
 |||||
 7312 CCGGTATTAACAGGTGGTGGCATCAATCTGGGTAAACCCCGCTGTG 7361
 417 lSerThrSerValPheAsnGlnSerLysAspLeuCysValMetValGlnI 434
 |||||
 7362 TTCACCTCATCTCTCAACCAATCCAAAGATCTCTGTCATGCTCCAAA 7411
 434 lValProArgValTyrTyrHisProGlnGlnValValLeuAspGlnTyr 450
 |||||
 7412 TCGCCCCCGAGTGTACTACCATCCGAGAGAGTGGTCTTGATGATAT 7461
 451 AspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAl 467
 |||||
 7462 GACTATGCTATTAACCCACCAAAAGAAACCCGTATCCCTTACCTAGC 7511
 467 aValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAlaA 484
 |||||
 7512 TGTAATGCTCGGATTAGGAGCGGCGCTTGGCGTAGAAGACAGGACAGCTG 7561
 484 lAlaLeuIleThrGlyProGlnGlnLeuGlnLysGlyLeuGlnIleHis 500
 |||||
 7562 CCCGATCACAGGACACAGAGCTAGAGAAAGACTTGGTAGAGCTACAT 7611
 501 AlaAlaMetThrGlnAspLeuArgAlaLeuLysGlnSerValSerAsnLe 517
 |||||
 7612 GCGGCGCATGACAGAGATCTCCGAGCTTAGAGAGAGCTGTGAGCAACT 7661
 517 uGlnGlnSerLeuThrSerLeuSerGlnValValLeuGlnAsnArgArg 534
 |||||
 7662 AGAAGAGTCCCTGACTTCTTCTGAAGTGGTCTACAGAACCCGAGAGG 7711
 534 lLysLeuAspLeuLeuPheLeuArgGlnGlyGlyLeuCysAlaAlaLeuLys 550
 |||||
 7712 GATTAGATCTGCTGTTCTTAAGAGAGAGTGGTTATGTGACAGCCTTAA 7761

551 GlnGlnCysCysPheTyrValAspHisSerGlyValAlaIleArgAspSer 567
 |||||
 7762 GAAGATGTTCTTCTTATGTATGATCCTCAGAGACCATCAGAGACTCAT 7811
 567 lAsnLysLeuArgLysLysLeuGlnArgArgArgGlnArgGlnAlaA 584
 |||||
 7812 GAGCAGACTTAGAGAAAGTTAGAGAGCGCTCGAAGGAGAAAGAGAGCTG 7861
 584 spGlnGlyTrpPheGlnGlyTrpPheAsnArgSerProTrpMetThrThr 600
 |||||
 7862 ACCAGGGGTGGTTTGAAGAGATGTTCAACAGCTCTCTTGATATACACC 7911
 601 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeuTh 617
 |||||
 7912 CTGCTTTCCTGCTGACGGGGCCCTTACTAGTCCGCTCGCTTACTTAC 7961
 617 rValGlyProCysLeuIleAsnArgPheValAlaPheValArgGlnArg 634
 |||||
 7962 AGTTGGGCGCTTGAATTAATAGCTTGTGCTTGTAGAGAACGAG 8011
 634 aLysAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeuLeu 650
 |||||
 8012 TGAGTGCAGTCCAGATCATGCTACTTAGGCAACACTACCAAGGCTTCTG 8061
 651 SerGlnGlyGlnThrAspLeu 657
 |||||
 8062 AGCCAAGAGAAACTGACCTC 8082
 seq_name: gb_v1:PERY17013
 seq_documentation_block:
 LOCUS PERY17013 7808 bp DNA VRL 14-APR-2000
 DEFINITION Porcine endogenous retrovirus gag, pol, env genes.
 ACCESSION Y17013.1 GI:3005985
 VERSION Y17013.1 GI:3005985
 KEYWORDS env gene; gag gene; pol gene.
 SOURCE porcine endogenous retrovirus.
 ORGANISM Viluses: Retroviral viruses; Retroviridae; Mammalian type C retroviruses: 1-Mammalian type C virus group.
 REFERENCE 1 (bases 1 to 7808)
 AUTHORS Gauderns, F., Fischer, N., Boller, K., Kurth, R. and Tonjes, R.R.
 TITLE Establishment and characterization of molecular clones of porcine endogenous retroviruses replicating on human cells
 JOURNAL J. Virol. 74 (9), 4028-4038 (2000)
 MEDLINE 20219394
 REFERENCE 2 (bases 1 to 7808)
 AUTHORS Tonjes, R.R.
 TITLE Direct Submission
 JOURNAL Submitted (30-MAR-1998) R.R. Tonjes, Paul-Ehrlich Institut, Paul-Ehrlich Str. 51-59, P.O. Box 1740, D-63207 Langen, FRG
 FEATURES
 source
 1..7808
 /organism="porcine endogenous retrovirus"
 /db_xref="taxon:61673"
 /tissue_type="kidney"
 /cell_line="PK15"
 /gene="gag"
 /codon_start=1
 /protein_id="CAA76581.1"
 /db_xref="GI:3005985"
 /db_xref="SnpREMBL:073504"
 /translation="MSGTVPPLSTITDHTWEVRSRAHNLVYKKKPMOTFCASEMP
 TEDVGMPSEGFNSFILLAKKAITFOIGRSHSDQERYITWDLIAEDPPMPKPMIN
 KPRAPGRILLALGKKNHSAKEVPSRITPEIEEPPTPEPPVPPPPVPPAGAVRG
 PSAPGAPVVEGPAAGTGRSRGATPERTDEIALILPLTVGPMGGLDLPQVPPSS
 ADLVNMTNIHPESEDPORLGLVSEIMPSHOPWDCODLQTLFTEERERILLEA
 RKNVAGDGRTOLEIDMGFLTRIGMDYNTAEGRESLKIYRQALVAGLRASRRP
 TNLAKREVWGQPNPEPSVFLERIMEAFRRFPDPPTSQAOKASVALAFIGOSALDIR
 KILQRLKGLDAELRDILVREAEKAYIGRETBEEKKEREERERDRRDRKKNIT

gene

KILAAVVEGKSSRRERDFRKISGPRQSGNLGNRPFLDKQCAVCEKEMWANCNK
KGNKPVLALEEDKD*

CDS

2120..4726
/gene="pol"
2120..4726
/gene="pol"
/codon_start=1
/protein_id="CA76582.1"
/db_xref="GI:3005987"
/db_xref="SPTREMBL:073505"
/translation="MGAOISFEQGRPEVSVNNKPITVLTLOLIDEXRIVSPQKPPD
IOSWLEQFQMAETAGMLAKOVPPVIOLKASATPVSRQYPLDSREAREGIVPVQ
RLIOGGLIVYOSPMNPLPRKPGINDRPVODLREVKRVODIHPIVNPYNLLS
ALPERKMYVLDLDAFELRLHPTSPQLAEEMRDPGRTGOLWTLRDPGFNS
PTJFEDALHRLANFRLOHPQVTLQYVDLLAGATKOCBLSTALLLEISDLGR
ASAKQICREYVLYGLSLRGGRMLTEARKTVVOIPAPYAKOVREFIGFGR
LMAJGFATLAAPLYPLTEKEGSEMAPEHOKADAKKALLSAPALADYKPPPLY
LQONITYAHALENIYKPPDKMTNARHTHOSLTERVTFAPPALINPATLLE
ETDEPVTHDCHLLIETGVKDLTDIPLTEVLTWETDGSYVVGKRMAGAAVVG
THTWASSLEPGTSAQALMALTOALRLAEGSINIYDSRYAFATVHGAIVKOR
GLTASGREIKNEELISLEALHLEKRLAIHCPGHOAKDLISNGMDADVAKOA
AOAVNLPIIETRAPPEPRQVLTEDMOETKKIDQSETPEGCTYSYKELIPHEG
LEVVOOIHRLTHTKHLQOLVTPSPYHVLREGVADSVYKHCVPQOLVNNAPSRILP
GKLRGSHBGAWHEVDPTEKYPKYNKTLVFDVDSGWEAYPRKKESTVAVAKY
WRKFRLDELTL*

gene

5294..7267
/gene="env"
5294..7267
/gene="env"
/codon_start=1
/protein_id="CA76583.1"
/db_xref="GI:3005988"
/db_xref="SPTREMBL:073506"
/translation="MHPILSWRLHPTRGVNGSTRGVAPRGTWMPPLHICRLINAVKSTP
RLIDSSNPHRPLSLTWLIDPHTGVNSTRGVAPRGTWMPPLHICRLINAVKSTP
PMLVRSYFGCCGTGTEKEREKCGSGSESPCRNRSCVTSNDGDKWPISTDRKESVN
SGRGKRYKMKLYKDKSCSPSDIDYLTISTFTEGKOINIDKWMNGMSWIVFYKYGGA
GSTLJLRLEIETGTEPPVAVGPDVYLAEGPAPLEPHNLVPOLTLSPDITOPPSN
GTTGLIPLTPRNSPGVPYKQRLSLIOGAFOALINSTDPATSCWLCSSGPY
EGMAKSGFNVTKEHNRQCTGMSRNKLTLEYSAGTCTGKAPSPHOLHCTSYVVEO
ASBNQVLVGVNMMACNTGLTPCVSSSVNPOPCVWOJLVPRVYHHPERVVIDEY
DYRNRKREPVSITLAVMLGCTANGVGTCAALITGQOLKEKIGELHAAHTEDLR
ALEESYNSLESISTSEVYLVONRRGLDLPLREGLCALALECCCFVDSHGALRDS
MSKRLERLRRRRERADGWEFGWPNRSPMTTLLSALTGLVLLTLITVGPCLIN
REVALEFREVGVQIOMLVLMROQYQGLSGOETDL"

CDS

5294..7267
/gene="env"
5294..7267
/gene="env"
/codon_start=1
/protein_id="CA76583.1"
/db_xref="GI:3005988"
/db_xref="SPTREMBL:073506"
/translation="MHPILSWRLHPTRGVNGSTRGVAPRGTWMPPLHICRLINAVKSTP
RLIDSSNPHRPLSLTWLIDPHTGVNSTRGVAPRGTWMPPLHICRLINAVKSTP
PMLVRSYFGCCGTGTEKEREKCGSGSESPCRNRSCVTSNDGDKWPISTDRKESVN
SGRGKRYKMKLYKDKSCSPSDIDYLTISTFTEGKOINIDKWMNGMSWIVFYKYGGA
GSTLJLRLEIETGTEPPVAVGPDVYLAEGPAPLEPHNLVPOLTLSPDITOPPSN
GTTGLIPLTPRNSPGVPYKQRLSLIOGAFOALINSTDPATSCWLCSSGPY
EGMAKSGFNVTKEHNRQCTGMSRNKLTLEYSAGTCTGKAPSPHOLHCTSYVVEO
ASBNQVLVGVNMMACNTGLTPCVSSSVNPOPCVWOJLVPRVYHHPERVVIDEY
DYRNRKREPVSITLAVMLGCTANGVGTCAALITGQOLKEKIGELHAAHTEDLR
ALEESYNSLESISTSEVYLVONRRGLDLPLREGLCALALECCCFVDSHGALRDS
MSKRLERLRRRRERADGWEFGWPNRSPMTTLLSALTGLVLLTLITVGPCLIN
REVALEFREVGVQIOMLVLMROQYQGLSGOETDL"

BASE COUNT

2077 a 1971 c 2047 g 1713 t
ORIGIN

alignment_scores:

Quality: 3468.00 Length: 657
Ratio: 5.303 Gaps: 0
Percent Similarity: 99.543 Percent Identity: 98.021

alignment_block:

US-09-171-553b-10 x PERY17013

Align seg 1/1 to: PERY17013 from: 1 to: 7808

1 MethisProthleuSerTrpArgHisleuProthtrArgIlygIupr 17
|||||
5294 ATGCATCCACGTTAAAGCTGAGCTCCGAGCTCGGGGTGAGAGGCC 5443
17 OlyArgLeuArgIleProleuSerPheAlaSerIleAlaTrpPheLeu 34
|||||
5344 GAAAGACTGAGTCCCTTAAGCTCGGCTGCATCGCTGCTGCTT 5393
34 hileuThrlleThProGlnAlaSerSerIysArgleuIleAspSerSer 50
|||||
5394 CTCTAACAAATAACTCCCGAGGCAAGTAGTAAGCGCTTATAGACAGCTCG 5443
51 AspProHisArgProleuSerProthtrPheIleIleAspProAspm 67

5444 AACCCCATAGACCTTATCCCTTACCTGGCTGATTTATGACCCCTGATAC 5493
67 rGlyAlaThrValAsnSerThrArgIlyValAlaProArgIlyThrTrp 84
5494 GGGTGTACTGTGTAATAGCACTCGAGGTGTTCCTCAAGAGGCCACCTG 5543
84 rPProGluLeuHisPheCysLeuArgleuIleAsnProAlaValIlySer 100
5544 GGCCTGAACCTGATTTCTGCTCCGATTTGATTAACCCCGCTTAAAGC 5593
101 ThrProAsnLeuValArgSerTrpGlyPheTrpCysCysProGlyTh 117
5594 ACACCTCCCAACCTAGTCCGAGTATGGGTGATATTGCTGCCAGCCAC 5643
117 rGluIlyGluIlyTrpCysGlyIlySerGlyIlySerPheCysArgArg 134
5644 AGAGAAAGAGAAATCTGTGGGGGTTCTGAGGAATCTTCTGTAGAGAT 5693
134 rPserCysValThrSerAsnAspIlyAspTrpIlySerProIleSerLeu 150
5694 GGACCTGCTGCTACCTCCACAGATGAGACTGGAATGGCCGATCTCTC 5743
151 GlnAspArgValIlyPheSerPheValAsnSerGlyProGlyIlyTrp 167
5744 CAGACCGGGTAAATATCTCTTGTCAATTCGGCCGGGCAAGTACAA 5793
167 smetMetIlyLeuTrpIlyAspIlySerCysSerProSerAspLeuAsp 184
5794 AGTAGTAACACTATATATAAGATAGAGCTGCTCCCATCAGACTTGAAT 5843
184 rIleuIlyIleSerPheThrGluIlyGlyIlyGlnIlyAsnIleGlnIly 200
5844 ATCTAAAGATAGATTCTCAGTAAAGAAAGAAAGAAATATTCAAAG 5893
201 TrpIleAsnGlyMetSerTrpGlyIleValPheTrpIlyTrpIlyGly 217
5894 TGGATTAATGTATGAGCTGGGGAATAGTTTATTAATAATAGCGGGG 5943
217 yAlaIlySerThrLeuThrlleArgleuArgIleGluThrArgIlyTh 234
5944 AGCAGGCTCCTTAACTTCCCTTGGAGTAAAGAGAGGAGGAGCAAGC 5993
234 rProValAlaMetGlyProAspIlyValIleAlaGluIlyGlnIlyPro 250
5994 CCCCTGGAGAGTGGGACCCGATTAAGTACTGCTGAACAGGGGCCCGG 6043
251 AlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeu 267
6044 GCCCTGAGGACACCCGATTAAGTCTGCGGCGCCCAATTAACCTCGCTG 6093
267 rProAspIleThrGlnProProSerAsnSerThrArgIlyLeuIlePro 284
6094 GCTTGACATTAACACACCCGCTAGCAGAGGTACCACTGATTAATGAT 6143
284 hTrAsnThrProArgAsnSerProGlyValProValIlyThrGlyGln 300
6144 CCAACAGCGCTAGAAACCTCCAGAGTTCCTGTTAAGCAGAGACAGAGA 6193
301 LeuPheSerLeuIleGlnIlyAlaPheGlnAlaIleAsnSerThrAsp 317
6194 CTCTTACGCTCATCCAGGAGCTTCCCAAGCATCACTCAGCAGACC 6243
317 rAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTrp 334
6244 TGAATGCACTTCTTCTGTTGGCTTGTCTATCCACAGGCTCTCTTAT 6293
334 rGluGlyMetAlaIlyGluArgIlyPheAsnValThrIlyGlnIlyArg 350
6294 ATGAGGGAGTGGCTAAAGAGAAATTCATAGTGCACCAAGAGCATAGA 6343
351 AsnGlnCysThrTrpGlySerArgAsnIlyLeuThrlleGluValIle 367

6344 AATCATATGACATGGGGGTCGCCGAATAAGCTTACCTCCTACTGAAAGTTTC 6393
 367 TGLYSGLYTHRCYSIIIEGLYLSALAPROPROSERHISGLNHISLEUC 384
 6394 CGGGAAGGAGCATGATAGGAAAGCTCCCAATCCACCAACACACTTT 6443
 384 YSTYSETHRVAlValTyrGlnAlaSerGluAsnGlnTyrLeuVal 400
 6444 GCTATAGTACTGTGTTATGACGAGGCTTCAAGAAATCAGTATTAGTA 6493
 401 PROGLYTYRASNARGTRTPALACYSANTRHGLYLEUTHRPROCYSA 417
 6494 CTTGTTTAAACAGTGGTGGGATGATCACTACTGGCTTAACCCCTGTGT 6543
 417 IserTherSerValPheAsnGlnSerLysAspLeuValMetValGlnI 434
 6544 TTCCTCAGCTCTTCAACCAATCCAAAGATTCTGTCTCATGGCCAAA 6593
 434 LeValProArgValTyrTyrHisProGlnGluValLeuAspGluTyr 450
 6594 TCGTCCCGAGTGTACTACCTCCTGAGGAAGTGTCTCTGATGAAATAT 6643
 451 ASPTYRATGYRASNARGPROLYSARGGLUPROVALSERLEUTHRLEUAI 467
 6644 GACTATCGGTATACCGACCAAAAGAGACCCGATCCCTTACCTAGC 6693
 467 eValMetLeuGlnGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAla 484
 6694 TGTATATGTCGAGATTAGGAGCGCGCTGGCTAGGACAGAGCAGCTG 6743
 484 lAlaLeuIleThrGlyProGlnGlnLeuGlnGlyLeuGlnGlnLeuHis 500
 6744 CCGTATATCACAGGACACAGCAGCTAGAGAAAGACTGGTGAAGCTACAT 6793
 501 AlaAlaMetThrGluAspLeuArgAlaLeuLysGlnSerValSerAspLe 517
 6794 GCGGCCATACAGAGATCTCCGACCTTAGAGAGTCTGTTAGCAACCT 6843
 517 uGlnGlnSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 534
 6844 AGAAGAGTCCCTGACTTCTTGTGAAAGTGTCTACAGAACCGAGG 6893
 534 lYLeuAspLeuLeuPheLeuArgGlnGlyGlyLeuCysAlaAlaLeuLys 550
 6894 GATTATATCTGCTGTTCTTAAAGAGAGGCTGATATGTCACGCTTAAA 6943
 551 GlnGlnCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMe 567
 6944 GAAGATGTTGCTTCTATATGATCACTAGAGACCATCAGAGACTCCAT 6993
 567 tAsnLysLeuArgLysLysLeuGlnArgArgArgArgGluAlaAla 584
 6994 GAGCAAGCTTAGAAGAGGTAGAGAGCGCTCGAAGGAGAAAGAGCGTG 7043
 584 spGlnGlyTyrPheGlnGlyTyrPheAsnArgSerProTyrMetThrThr 600
 7044 ACCAGGGGTGTTTGAAGATGTTTCAACAGGCTCTCTTGATGACACAC 7093
 601 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuThr 617
 7094 CTGCTTTCGCTCTGAGGGACCCCTAGTAGTCTGCTCTTACTTAC 7143
 617 rValGlyProCysLeuIleAsnArgPheValAlaPheValAlaGluValArg 634
 7144 AGTTGGGCTTGTCTTAAATTAATAGGTTTGTGCTTGTGTTAGAACGAG 7193
 634 alSerAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeuLeu 650
 7194 TGGGTCCAGTCCAGATCATGTAAGTACCAACAGTACCAAGGCCCTTCTG 7243
 651 SerGlnGlyLysThrAspLeu 657
 7244 AGCCAAAGGAGAACTGACCTC 7264

seq_name: gp_v1:PEN133818

seq_documentation_block: 8750 bp DNA VRL 14-Apr-2000

LOCUS PEN133818 Porcine endogenous retrovirus type C proviral gag, pol and env genes and LTR (class B, clone 43).

DEFINITION genes and LTR (class B, clone 43).

ACCESSION AJ133818

VERSION AJ133818.1 GI:6688949

KEYWORDS env gene; env protein; gag gene; gag protein; pol gene; pol protein.

SOURCE porcine endogenous retrovirus.

ORGANISM Porcine endogenous retrovirus.

REFERENCE 1 (bases 1 to 8750)

AUTHORS Czauderna, F., Fischer, N., Boller, K., Kurth, R. and Tonjes, R.R.

TITLE Establishment and characterization of molecular clones of porcine endogenous retroviruses replicating on human cells

JOURNAL J. Virol. 74 (9), 4028-4038 (2000)

MEDLINE 20219394

REFERENCE 2 (bases 1 to 8750)

AUTHORS Toenjes, R.R.

TITLE Direct Submission

JOURNAL Submitted (04-MAY-1999) Toenjes R.R., Medical Biotechnology, Paul-Enrich-Institut, Paul-Enrich-Strasse 51-59, Hessen, D-63225 Langen, GERMANY

FEATURES

source

Location/Qualifiers

1..8750

/organism="porcine endogenous retrovirus"

/proviral

/strain="type C"

/db_xref="taxon:61673"

/clone="43"

1..629

467

/note="cap site"

632..649

/note="gag"

1064..2635

/gene="gag"

1064..2635

/gene="gag"

/function="retroviral capsid"

/product="gag protein"

join(1064..2635,2639..6223,6099..8072)

/note="viral polypeptide"

/codon_start=1

/protein_id="G1:6688950"

/db_xref="GI:6688950"

/translation="MGCTVTTPLSLTLDHMTVRSRAHNLSTOVYKKGPMOTFCASEMP

TFDVGMPSEGTNSSTILAVKALIFOTGSSHDDEPYLTLMODIADPPVAKFWLN

KRRKPGRIALAGEKNKHSKAEVPEPSRTPYPIEBPMPPEPPVAPGAVFG

PSAPGAPVVEGPAAGTSGRSRATPERDEIALILPRYGPMPGOLQPLQWPFSS

ADLYNKKTNHPPESDPORLTGLVESIMSHQPTMDCCOLLOTLETERERILLEA

RKNVREAGDSRPOLONEIDMGPRITRPGMDYTACGRESCTKYRQALVAGASRRP

TNLAKVREVMGPNPEPSVLERLMEARTRTPPTDSQAQASVALAFIGSALDIR

KLQRLGELQEAELRDVLEAEKVVYRRRETEKQRKEREERERDRKQENLT

KLLAAVEGKSSREEREDRKRISGRSGNNGNTPLDKDQACAKCKEGRHAKNCPK

KGNKGKVLALDEEDKGRGSDPLEPRVALVEGQPVFLVDGAESHVLLQPLGKL

KKKSWMGATGEGORVPMTTRTVDLAVGRVHSFLVPECPVLLGRLTKMGAOI

SPEQGRPEYVNNKPTTVTLTLDLDEYRLYSPQVPPDOIOSMLDOPVMAETGAG

LAKOVPPVIOIKASATPVSVQYPLSRARSGIMPHVORLIQOGLVLPVDSFMTPL

LVRRKPGTIDRPVODLREVNKRVDIHTVNPYNLISALPERNWTVDLKAATF

CLRIHPTSQLFAEWRDGTGRTGOLWTRLDQFKNSPTIFEDALHDLANFRIQH

POVTLLOYDILLAGATKODCIETGKALLILSLDGLYASAKKAQICREVTYLGK

LGGGSMILTEARKTIVYVPIPTTAKOVREPLGTAGFCRLAIPGATLAAPLPYLKE

KGEFASHPHQCTPAIRKALLSAPALADPVTKPTTYLVDKRGVAGVLTOTLGPW

KRPVALISKKLPDVGSGVCLKALNAVALYVKDADKLTIGNTIVYAPHAENIVRO

PPDRMTNARNMTHYOSLLTERVTEAPPAALPATLLPEETDEVTYHDCOLLIEG

VRKDLTDLPTGEVLITWFTDSSYVEGKRMAGAVVDGTHITASSLSPEGTSQAKAE

LMALTOALAEKGSINITYDSRYAFATNHAIVGAIYKQGLTSGREIKNKEELISL

LEALHLPKRLALIIHCPGHOKARDLISRGNOMADRVAKQAQAVNLLPIIEPKADEPR

RQYLTEDMDEIKKIDOFSEPECTCYTSYCKELLPHKEGLEEVYQOILHRLHLETKHLO
 QLVATSPYHVLRLPGVADSVYKHCVPOLVNANPSPGKRLRGSHPAGMVDTE
 VKPAKYNNKLLVVDPEVSGMVEAYPTKKTSTVYAKKILEELFEPFGIPKVGMS
 PAFVAYOSGLAKILGIDMKLHCAYRPOSSGVOYRNNKTKETLTETGNDMTA
 LLPVLEFRVNTPGGORTIPYELLGSPPLVIAVSHADVLSLSPLSRLKALMAY
 KORAMKOLREAYSEEDLOVPHRFQVDSTVYKRRHAGNLETWKGPYLVLLTPYAV
 KVEGISTWISHAVKLAPPEPDSGMRAKTEENPKLRLHRLYRNNNSGOMPLISM
 RHLPTRGGERKLRLPLSFASIAMFLITTPQASKRLLDSNPHRPLSLTWLLIDP
 DTGVTWSTRGVAPRGTMWDELHCLRLINPAKSTPNLVRSGYCCPGTEKEXC
 GGSSEFCRMSCVTSNDGMKWPISLODVEKSFVNSGPKYKWKLYKDKSCSPSD
 LDYLTISFEKGOENIKNMGSMGIVFYKGGAGSTLTLRLTETGTEPPAVG
 PDKVLAEQGPALPEPHNLPVPLTSLRPDITDPPSNGTTLPLTPNPSGVPVKT
 GORFLSLIQAFQAINSTDPAVSCWLCISGPPYEGNAEKGNVTKENHNOCTM
 GSRNKLTLFVSGKTCIGKAPPSHQHLCYSTVYVQASENOVLPGYRMMACNTGL
 TPCVSSSVFNOSKDFCVMOIYPRVYHPEEVYVDEYDRYRNPKRREPSLTILAVLG
 LGTAVGVTGTLALITGPOOLEKELGELAEAMEDRLAESYLNESLTSISVYL
 ONRRGDLFLREGGLCALKECCCFYVDHSGAIRMSPKSRRLERRRRERKADGW
 FKGMFNRSPMTLLLSNLGPLVLLLLTVGCLINRFVAFREKVSAYQIMVLKQO
 YQGLSOGERTDL
 2639 .6223
 /gene="pol"
 /function="retroviral protease and polymerase"
 /product="pol protein"
 2639 .6223
 /gene="pol"
 /product="pol protein"
 6099 .8069
 /gene="env"
 /product="retroviral envelope"
 /product="Env protein"
 6099 .8069
 /gene="env"
 /product="Env protein"
 8122 .8750
 8643 .8648
 2214 c 2286 g 1962 t
 BASE COUNT 2288 a 2214 c 2286 g 1962 t
 ORIGIN
 LTR
 polyA_signal
 8643 .8648
 2214 c 2286 g 1962 t
 BASE COUNT 2288 a 2214 c 2286 g 1962 t
 ORIGIN
 alignment_scores:
 Quality: 3468.00 Length: 657
 Ratio: 5.303 Gaps: 0
 Percent Similarity: 99.543 Percent Identity: 98.021
 alignment_block:
 US-09-171-553B-10 x PEN133818 ..
 Align seg 1/1 to: PEN133818 from: 1 to: 8750
 1 MethisprothleuSerTrrpArgHisleuProthrrarglygllylupr 17
 |||||||
 6099 ATGCATCCACGTTAAGCTGGCGCACCTCCCGACCTGGGGTGAGAGCC 6148
 |||||||
 17 olyAargleuAargileProleuSerPheAlaSerIleAlaTrpPheLeu 34
 |||||||
 6149 GAAAGACTGAGAAATCCCTTAACCTTCGCTCCATCGCCCTGGTTCCTTA 6198
 |||||||
 34 hrleuThrlleThrProGlnAlaSerSerlySargleuIleasPserSer 50
 |||||||
 6199 CTCTAACAAATTAATCTCCAGGCCAGTAGTAACCCCTTATAGACAGCTCG 6248
 |||||||
 51 AsnProHisArgProleuSerProThrTrpLeuIleIleasProAspTh 67
 |||||||
 6249 AACCCCATAGACCTTTATCTTACCTGGCTGATTTATGACCCCTGATAC 6298
 |||||||
 67 rGlyValThrValAsnSerThrArgGlyValAlaProArgIlyThrTrp 84
 |||||||
 6299 GGGTGTCACTGATAATAGCACTGAGGTGTGGCTCCTAGAGGACACCTGGT 6348
 |||||||
 84 rPrProGlnLeuHisPheCysLeuAargleuIleAsnProAlaValIysSer 100
 |||||||
 6349 GGGCTGAACTGCATTTCTGCCCTCGATTAACCCCGCTGTAAAGC 6398
 |||||||
 101 ThrProAsnLeuValAargSerTyrGlyPheTyrCysCysProGlyTh 117
 |||||||
 6399 ACACCTCCCAACCACTAGTCGAGTATAGGCTCATATGCTGCCACGAC 6448
 |||||||
 117 rGlyValGlnIlyThrTyrCysGlySerglyGlnSerPheCysArgArg 134
 |||||||
 6449 AGAAGAAAGAAATTAATCTGGGGGTCTGAGGAATCTTCTGTAGAGAT 6498
 |||||||
 134 rPserCysValThrSerAsnAspGlyAsnPrPlySTppPoiIleSerLeu 150
 |||||||
 6499 GGACCTCGTCACCTCCACAGATGAGACACGAAATGCCGATCTCTCTC 6548
 |||||||
 151 GlnAspArgValIlySerpheSerPheValAsnSerglyProGlyIlyStry 167
 |||||||
 6549 CAGAGCCGGGTAAATCTCTCTTGTCAATTCGGCCCGGCAAGTAACA 6598
 |||||||
 167 smelMelIlyLeuThrIlySaspIlySerpCysSerProSerAspLeuAsp 184
 |||||||
 6599 AGGATTAACACTATATTAAGATTAAGACCTCTCCATCGACTGTAGATT 6648
 |||||||
 184 yrlleuLysIleSerPheThrGlnIlyGlyIlySglnIlyAsnIleGlnIly 200
 |||||||
 6649 ATCTAAAGATTAAGTTTCACTGAAAAAGAAAAAGAAAAATATTCAAAA 6698
 |||||||
 201 TrpIleAsnGlyMetSerTrpGlyIleValPheTyrIlyStryGlyIyl 217
 |||||||
 6699 TCGATTAATGTATAGCTGGGGAATAGTTTTTATTAATATGCGGGGG 6748
 |||||||
 217 yAlaGlySerThrLeuThrIleAargleuAargileuIlyThrGlyThrG 234
 |||||||
 6749 AGCAGGGTCCACTTTTAACCACTTCGCTTAGAGTAGAGCGGGGAGAAC 6798
 |||||||
 234 roProValAlaMetGlyProAspIlyValIleuAlaGlnIlyPropo 250
 |||||||
 6799 CCCCTGGCAGTAGGGAGCCCGATTAAGTACTGCTGACACAGGGGCCCG 6848
 |||||||
 251 AlaLeuGlnUProProHIsanLeuProValProGlnLeuThrSerLeu 267
 |||||||
 6849 GCCCTGGAGCCAGCGATTAATCTTCCGCTCCCAATTAACCTCGCTGG 6898
 |||||||
 267 gProAspIleThrGlnProProSerAsnSerThrTrpGlyLeuIlePro 284
 |||||||
 6899 GCCTGACATTAACACAGCCGCTTACGACACGGTACCTGATGATTCCTA 6948
 |||||||
 284 hrAsnThrProArgAsnSerProGlyValProValIlyThrGlnIly 300
 |||||||
 6949 CCAACAGCGCTAGAAATCCCGAGGTCTTCGTTTAAGACAGACAGAGA 6998
 |||||||
 301 LeuPheSerLeuIleGlnIlyAlaPheGlnAlaIleAsnSerThrAsp 317
 |||||||
 6999 CTCTTACAGTCTATCCAGGGAGCTTTCAGGCCATCACTCCACCGACC 7048
 |||||||
 317 oAspAlaThrSerSerCysTrpLeuCysLeuSerSerglyPropoTyr 334
 |||||||
 7049 TGTATGCACCTTCTTGTGCTTGGCTTTGTTATCTTCAGGGCTCTCTAT 7098
 |||||||
 334 yrlGlnIlyMetAlaIlySglnAargIlySerpheAsnValThrIlySgln 350
 |||||||
 7099 ATGAGGGGATGGCTAAGAAAGAAATTCATATGTGACCAAGACAGATGA 7148
 |||||||
 351 AsnGlnCysThrTrpIlySerArgAsnIlySleuThrlleuThrGlnVal 367
 |||||||
 7149 AATCAATATGATCAATGGGGGTCCGAAATTAAGCTTACCTACAGAAAT 7198
 |||||||
 367 rGlyIlySglnThrCysIleGlyValAlaProProSerHisGlnHisLeu 384
 |||||||
 7199 CGGGAAGGGGACATGATAGAAAAAGCTCCCATCCCAACACACCTTT 7248
 |||||||
 384 yslIlySerThrValIlyrGlnIlyAlaSerGlnAsnGlnIlyLeuVal 400
 |||||||
 7249 GCTATATGATCTGGTTATGAGCAGGCTCGAAGAAATCAGATTTAGTA 7298
 |||||||
 401 ProGlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysVal 417
 |||||||
 7299 CCTGTTATACAGGTGGTGGCATGCAATACTGGTTAAACCCCTGTGT 7348
 |||||||

417 |SerThrSerValPheAsnGlnSerLysAspLeuCyValMetValGlnI 434
 7349 |TTCCTCCCTCAGCTTCAACCAATCCAAAGATTCTGTGTCATGGCTCAAA 7398
 434 |LeValProArgValTyrTyrHisProGluGluValValLeuAspGluYr 450
 7399 |TGTCTCCCGAGTGTACTACATCTCTGAGGAAGTGTCTCTGTGATGAATAT 7448
 451 |AspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAl 467
 7449 |GACTTCGCGGTATACCGAACCAAAAGAGACCCGTATCCCTTACCCTAC 7498
 467 |aValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAla 484
 7499 |TGTATATGCTCGGATTAGGAGCGCGCTTGCGCTAGAACAGGACAGCTG 7548
 484 |IaLeuLlIeThrGlyProGlnGlnLeuGlyLysGlyLeuGlyGluLeuHis 500
 7549 |CCCTATATACAGACACACAGCAGCTAGTGAAGAAAGACCTTGATGAGCTACAT 7598
 501 |AlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnLe 517
 7599 |CGGCGCATGACAGAAAGATCTCGAGCCTTAGAGAGTCTGTTAGCAACCT 7648
 517 |uGluGluSerLeuThrSerLeuSerGluValValLeuGluAsnArgArg 534
 7649 |AGAAAGATCCCTGACTCTTGTCTGAGAGTGCTTCTACAGAACCGGAGG 7698
 534 |LysLeuAspLeuLeuPheLeuArgGluGlyGlyLeuCyValAlaLeuLys 550
 7699 |GATTGATCTGCTGTTCTTAAGAGAGGTGGCTTATGTGACCCCTTAATA 7748
 551 |GluGluCysCysPheTyrValAspHisSerGlyAlaIleAspSerSeMe 567
 7749 |GAAGATGTTGCTTCTATGTAGTACTACTCAGAGCCATCAGAGCTCCAT 7798
 567 |TasulLysLeuArgLysLysLeuGluArgArgArgGluArgGluAla 584
 7799 |GAGCAAGCTTAGAGAAAGTTAGAGAGCGCTGCAAGAGGAAGAGAGCTG 7848
 584 |spGlnGlyTyrPheGlnGlyTyrPheAsnArgSerProTyrMetThrThr 600
 7849 |ACCAGGGGTGTAAAGATGTGTTCAACAGCTCTCTCTGATGACCAACC 7898
 601 |LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuThr 617
 7899 |CTGCTTCTGCTCTGACGGGACCCCTAGTACTCTGCTCTGTTACTTAC 7948
 617 |IValGlyProCysLeuLysAsnArgPheValAlaPheValArgLysArg 634
 7949 |AGTTGGGCTTGCTTAATTAAAGTTGTTGCTCTGTTGTTAGAACAGAG 7998
 634 |aLSeValValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeuLeu 650
 7999 |TGAAGCGAGTCAGATCATGTACTTAGCAACAGTACCAAGGCCCTCTG 8048
 651 |SerGlnGlyGluThrAspLeu 657
 8049 |AGCCAGAGCAAACTGACCTC 8069
 seq_name: gb_v1:PEM13816
 seq_documentation_block:
 LOCUS PEM13816 8918 bp DNA VRL 07-DEC-2000
 DEFINITION Porcine endogenous retrovirus type C proviral gag, pol and env
 genes and LTR (class B, clone 33).
 ACCESSION AJ133816
 VERSION AJ133816.1 GI:6688945
 KEYWORDS env gene; env protein; gag gene; gag protein; pol gene; pol
 protein.
 SOURCE porcine endogenous retrovirus.
 ORGANISM porcine endogenous retrovirus.

Viruses: Retroid viruses: Retroviridae; Mammalian type C
 retroviruses; 1-Mammalian type C virus group.
 REFERENCE
 1 (bases 1 to 8918)
 Czaderna, F., Fischer, N., Boller, K., Kuth, R. and Tonjes, R.R.
 Establishment and characterization of molecular clones of porcine
 endogenous retroviruses replicating on human cells
 J. Virol. 74 (9), 4028-4038 (2000)
 JOURNAL
 MEDLINE
 20219394
 REFERENCE
 2 (bases 1 to 8918)
 Toenjes, R.R.
 Direct Submission
 TITLE
 Submitted (04-MAY-1999) Toenjes R.R., Medical Biotechnology,
 Paul-Ehrlich-Institut, Paul-Ehrlich-Strasse 51-59, Hessen, D-63225
 Langen,
 GERMANY
 FEATURES
 Location/Qualifiers
 source
 1. 8918
 /organism="porcine endogenous retrovirus"
 /proviral
 /strain="type C"
 /db_xref="taxon:61673"
 /clone="33"
 /note="grown in PK15 (ATCC No. CCL-33) cell line"
 1..707
 545
 /note="cap site"
 710..727
 /note="gly4"
 1154..3725
 /gene="gag"
 1154..2725
 /function="gag"
 /product="retroviral capsid"
 join(1154..2725,2729..6313,6189..8162)
 /note="viral polypeptide"
 /codon_start=1
 /protein_id="CAB65339.1"
 /db_xref="GI:6688946"
 /db_xref="SPTREMBL:O9Q1X5"
 /translation="MGCTVTPPLSLTLDHMTFVRSRAHNLSVQYKKGPMQFCASEMP
 TEDVGMPESEGFNSITLAVKAILFQICSPSSHPPOEPIITLWDLAEDPPWVAVKPLN
 KPRKGRITLAGEKNKHSARKVESEPTIPEIEPPVPEPPVPPPPVPAQVAFGLN
 PSAPGAVVVEGPAAGTGRSRGATPERDEIALLPLRTYGPMMGGLOLQVDFSS
 ADLNMKATNHHPESDEPLRLGLVESLHESHPQWDCCOOLITLFTTEERILLBA
 RKNVGDAGRPLOLONEIDMGEPLTRPGMDVNTAAGRSALKITRQALVAGLRSSRP
 TNLAKVRVMGCPMPPEVSLELRMEARRRTPDPDSSEAKASVALAFQGSGLDRP
 KKLRLBGLQAEILDVAREAEKYIRRETEEEKQKEREERERERDRRQKSNLT
 KTLAAVEGKSSREERERDRIRSGPQSNLGNRTPLDKQCAVCKEKGHWARNCRK
 KGNKGPVLTLEBDKGRGSDPLPEBPVTLKVGQPEPEPLVDGASHSVLLQPLGL
 KEKSWMGATGQOYPTWTRRYDLAVGRTVHSFLVPECPVPLGLRDLITKMGAGL
 SEEGRPVSNKRPITVLTLDLDEDELYSPOYKPPQDIOSTLEOHPQMAETAGMG
 LAKOVPRVIOLKASATPVSVROYPLSREAREGIWPIVORIIOGGLIVPQSPNPTPL
 LPRKPGTNDIRPVODLREVNKRVQDHPVTPNYNLISALPEPRMYTVLDLDAEF
 CLRLHPTSPPLFAEWRDPTGRTGRLTWRLPDGFNNSPTIFEDALHRLDANRICH
 POVTLIOVVDLLLAGATKODCLEGTALLELSDLYRASAKAOIKREVEVLTGYS
 LRQGMRLTTEARKKTVVOIPAPTAQOUREPLTAGCCRLMIPGFATLAPVLTRE
 KGEFSMAPEHOKAPDAIKKALLSAPALAPVTPPEPLTYDEBKVGKVAAGVLTQISW
 RRPVAVLSKKTIQDPAVSGMPVCLKAIAVALIVKADKTLTGQNTVAPAHLEIVVO
 PPRKMTNARKTHVOSLITERVTFAPPAALNPVTLPEETDEVTIDCHQDLIEEG
 VRKDITDPLTGEVLWFTDSSYVVEGKRMAAAVVDGHTTWASSLPKTSQKKE
 LMAULTALRLAEKSLINVTDSRAVFAVAHGAIVKQGLIAGVSEIKNEKLEL
 LEALHLPKRLAIHQPGHOKAKDLISGNQMAADVAAQAAVNLPIITPKAPERR
 ROYLTLEDMOEIKIIDQSEPEGETCTSYGKEILPHEKGELEYOUIHRLHGTGKH
 OLVTSPYHVLRLPGVADSVYKHCVCPOLVANASRIIPGKRLRSGSGAHMEPPE
 VRPAKYGKYLIVVDFTSQVENVYPPKKESTYVAKTILEITPRGICRIVGSDNG
 PAFVAQVSQGLAKLIGIDMKLHCAYRQSSQVVRNMRITKETITKLTTEGIDMVA
 LILPEVLRVRNTPGQGLIPEVLLYGGPPVLEVAHSVADVLISQPLFRLKALEVY
 RORAMKOLREAVYSGEGLQVRFQVDSYVVRHRAAGNLETNRKGPYVLLTPTAV
 KVEIGSTVITSHVKLAPDPDGGWRAKTEPELKLRLHRLYVSNNSPOIIPITTSV
 RHLPDGEPEKRLRILPISFASIAWELITLTPPOKSRLLDSSNRRPLSLTNIIDP
 DTGYVNSTRCGVAPRGTMPELHICILINPAKSTPNLVRSRGPCCCTGTEKRC
 GDSGSEPRRMSCTVNDGDKRPTISLDQKRVKSFVSGGKTKVMKLVKDKSCSPSD
 LDYLIKSTFERKQENIOKWINGMSGIVFYKYGAGSTLITRLRIETGTETPPVAVG

```

PDKVLAEGPPALPEPPHNLVLPOLTSRLPDTIOSPSNGTGLIPTNTPRNSPGVEVKT
GDRFSLIOGAFOALINSDPATSSCMICLSGSPRYEGMAEKFNTKERNOCITW
GSRNKLTLFEVSGKGTGKAPPSHOHLCYSTVVYEOASENVYLVPYNNRMACNTGL
TPCVSTSVSPNOSKDRVMVOIYPRVYIHEPEVVDDEYIRYRPRKREVSILFLAMIG
LGTAVGVGTALITGPOLLEKGEJELHAAATELDRALESVSNLESSTLSLEVYL
QNRNGIDLLFLREGGICAALEKCECFYVDSGAIIDSMSKILNERLERRREREADOGW
FRGFNRSRPMPTTLLSALGPLVLLLLTVPCLINREFAVERRSVAVOIMVIRROO
YQILSOGSETD"
mat_peptide
2729..6313
/gene="pol"
/function="retroviral protease and polymerase"
/gene="pol"
/product="Pol protein"
/gene="pol"
2729..6313
/gene="env"
/function="retroviral envelope"
/product="Env protein"
/gene="env"
6189..8162
/gene="env"
8212..8918
2327 a 2245 c 2334 g 2012 t
BASE COUNT
ORIGIN
alignment_scores:
Quality: 3465.00 Length: 657
Ratio: 5.306 Gaps: 0
Percent Similarity: 99.391 Percent Identity: 98.021
alignment_block:
US-09-171-553B-10 x PEN133816 ..
Align seg 1/1 to: PEN133816 from: 1 to: 8918
1 MethisProthrLeuSerTrpArgHisLeuProThrArgGlyGlyLeu 17
:::|||||
6189 ATACATCCACGTTAACTGAGCTGGCGCACCTCCGACCTCGGSGTGAGAGACC 6238
17 OlysarGleuArgGlyLeuProLeuSerPheAlaSerIleAlaTrpPheLeu 34
|||||
6239 GAAAAGACTGAGAAATCCCTTAAGCTTCGCCATCCATCCGCTGCTCTTA 6288
34 hrLeuThrIleThrProGlnAlaSerSerIleAspArgLeuIleAspSer 50
|||||
6289 CTCTACAAATACCTCCCGAGGCCAGTAAACGCTTATAGACAGCTCG 6338
51 AsnProHisArgProLeuSerProThrTrpLeuIleIleAspProAspTh 67
|||||
6339 AACCCCAATAGACCTTTATCCCTTACCTGCGCTGATTTATGACCTGATAC 6388
67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84
|||||
6389 GGGTGTCACTGTAATAGACACTCGAGGTGTCTCCTAGAGGCACTCGT 6438
84 rPpGlnLeuHisPheCysLeuArgLeuIleAsnProAlaValIleAspSer 100
|||||
6439 GGCTGAAGTGCATTTCGCTCCGATTTGATTAAACCCGCTGTGTTAAAGC 6488
101 ThrProAsnLeuValArgSerTrpGlyPheTrpCysCysProGlyTh 117
6489 AACCTCCCAACCTAGCTGCTAGTATGCTGTTATGCTGCTGACAGGAC 6538
117 rGlyValGlyValGlyValGlyValGlyValGlyValGlyValGlyVal 134
6539 AGAGAAAGAGAAATCTGTTGGGATTCGGGGAATCCCTTCTGTGAGGAT 6588
134 rPpSerCysValThrSerAsnAspGlyAspTrpTrpTrpProIleSerLeu 150
6589 GGAGTGGCTGACCTCCCAACGATGAGACTGGAATGCGCATCTCTCTC 6638
151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTrpLys 167
|||||
6639 CAGGACCGGGTAAATTCCTCTTTCATTCGCGGCCGCGGAGACAA 6688
167 SMeTLeuLysLeuTrpLysAspLysSerCysSerProSerAspLeuAsp 184
6689 AGTGAAGAACTATATATAAGATAGAGCTGCTCCCATAGACTATGATT 6738
184 YrLeuLysIleSerPheThrGlyValGlyValGlyValGlyValGlyVal 200
6739 ATCTAAAGATTAAGTTCCACTGAGAAAAGAAACAGAAATATTCAAAAG 6788
201 TrpIleAsnGlyLeuSerTrpGlyIleValPheTrpLysTrpGlyVal 217
6789 TGGATTAATGATGATGAGCTGCGGAATAGTTTATTAATAATAGCGGGG 6838
217 YAlaGlySerThrLeuThrIleArgLeuArgIleGluTrpGlyThrGlu 234
6839 AGCAGGGTCCACTTAACTTACCTTCCTTACGATAGAGACGGGACAGAAC 6888
234 rProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProPro 250
6889 CCCCTGGCGAGTGGGACCCGATAAGTACTGGCTGACACAGGGGCCCGCC 6938
251 AlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuAr 267
6939 GCCCTGAGGACCGACCATTAAGTCCGCGCCCAATTAACCTCGCTCGC 6988
267 rProAspIleThrGlnProProSerAsnSerThrTrpGlyLeuIlePro 284
6989 GCGTGAATATACACAGTGGCTTACGACAGGTACACTGATGATTCCTTA 7038
7039 CCAACAGCGCTAGAACTCCCGAGGTGCTGCTTAAAGACAGACAGAGA 7088
301 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAsp 317
7089 CTCTTACGTCATCCAGGAGGCTTCCAGGCTTCAAGCATCACTCACACGCC 7138
317 oAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTrp 334
7139 TGATGCCACTTCTTCTTGTGCTTGTGCTTGTATCTTACGAGGCTCTTAT 7188
334 YrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArg 350
7189 ATGAGGGGATGGCTAAAGAAAGAAATTCATGTGACCAAGAGCTAGTA 7238
351 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSe 367
7239 AATCAATGTACATGAGGGGTCCCGAAATTAAGCTTACCTCAGTCAACTTTC 7288
367 rGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeu 384
7289 CGGGAAGGGACATGATGGAAGAGCTCCCGCATCCACCAACACACTTT 7338
384 YrTrpSerThrValIleGlyGluGlnAlaSerGlnAsnGlnTrpLeuVal 400
7339 GCTATATATGTTGTTTATGAGCAGGCCCAAGAAATCAGATTTAGTA 7388
401 ProGlyTrpAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysVa 417
7389 CCGTGTATTAACAGGTGCGGATCAGTCAATACGAGGCTTAAACCCCTGT 7438
417 LserThrSerValPheAsnGlnSerLysAspLysValMetValGlnI 434
7439 TTCACCTCAGTCTTCAACCAATCCAAAGATTTCTGTGATAGTCCAAA 7488
434 LeValProArgValTrpTrpHisProGluGluValAlaLeuAspGlyTr 450
7489 TCGTCCCGCAGGTACTACCATCCGAGGAAGTGTCTTATGATGAATAT 7538
451 AspTrpArgTrpAsnArgProLysArgGluProValSerLeuThrLeuAl 467
|||||

```

7539 GACTATCGGTATACCGAACAAAGAGACCCGTATCCCTTACCCCTAGC 7588
 467 aValmetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAla 484
 7589 TGTATGCTCGGATTAGGACGCGCGCTGGCTAGAACACGACACCTCG 7638
 484 lAlaLeuIleThrGlyProGlnGlnLeuGlyLeuGlyLeuGlyLeuGly 500
 7639 CCTGATACACGACACGACGACGTAGAGAAAGACTGTGTGAGCTACAT 7688
 501 AlaAlaMetThrGluAspLeuArgAlaLeuGlyGluSerValSerAsnLe 517
 7689 GGGCCCATGACAGAGATCTCGACCTTAGAGAGTCTGTAGCAACCT 7738
 517 uGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 534
 7739 ACAAGAGTCCCTGACTCTTTGTGAGAGTGCTTACAGAACCCGAGG 7788
 534 lLeuAspLeuLeuPheLeuArgGluGlyLeuCysAlaAlaLeuLeuLys 550
 7789 GATTAGATCTGCTGTTCTTAGAGAGAGTGGTTATGTGCACCTTAA 7838
 551 GluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMe 567
 7839 GAAGATGTTGCTCTATGTAGATCATCTCAGACCATCAGACTCCAT 7888
 567 tAsnLysLeuArgGlyLysLeuGluArgArgArgGluArgGluAla 584
 7889 GAGCAAGCTTGAAGAAAGTTAGAGAGCGCTGGAAGGAAAGAGCGCTG 7938
 584 sPGLnglyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThrTr 600
 7939 ACCAGGGGTGCTTGAAGATGTTCAACAGSTCCCTTGATGACACAC 7988
 601 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeu 617
 7989 CTGCTTTCCTCTGCTGACGGACCCCTAGTAGCTGCTCTGTACTTAC 8038
 617 tValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArg 634
 8039 AGTTGGCCCTGCTTAATTAATAGGTTGTTGCCCTTGTAGAGACGAG 8088
 634 aLSerAlaValGlnIleMetValLeuArgGlnGlnIleGlnGlnLeu 650
 8089 TAGTGCATCTCCAGATCATGTACTTAGCAACAGTACCAAGCCCTTCG 8138
 651 SerGlnGlyGluThrAspLeu 657
 8139 AGCCAAGAGAAACGTGACCTC 8159
 seq_name: gb_pat:A66552

seq_documentation_block:
 LOCUS A66552 8196 bp DNA PAT 29-MAR-1999
 DEFINITION Sequence 2 from Patent WO9740167.
 ACCESSION A66552
 VERSION A66552.1 GI:4538105
 KEYWORDS
 SOURCE
 ORGANISM
 porcine endogenous retrovirus.
 Viruses; Retroviral viruses; Retroviridae; Mammalian type C
 retroviruses; 1-Mammalian type C virus group.
 1 (bases 1 to 8196)
 Galbraith,D.N., Haworth,C., Lees,G.M. and Smith,K.T.
 PORCINE RETROVIRUS
 Patent: WO 9740167-A 2 30-OCT-1997;
 JOURNAL
 Q ONE BIOTECH LTD (GB)
 location/Qualifiers
 FEATURES
 1..8196
 /organism="porcine endogenous retrovirus"
 /db_xref="taxon:61673"
 BASE COUNT 2165 a 2061 c 2146 g 1820 t 4 others
 ORIGIN

alignment_scores:
 Quality: 3445.50 Length: 658
 Ratio: 5.276 Gaps: 2
 Percent Similarity: 99.240 Percent Identity: 98.936

alignment_block:
 US-09-171-553B-10 x A66552 ..

Align seq 1/1 to: A66552 from: 1 to: 8196

1 MethisProThrLeuSerTrpArgHisLeuProThrArgGlyGlyLupr 17
 5606 ATGCATGCCACGTTAAACGCGGCACCTCCGACCTCGGGGTGGAGAGCC 5655
 17 oLysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeu 34
 5656 GAAAGACTGAGANTCCCTTAAAGCTTCGCTCCATCGCTGCTTCCTTA 5705
 34 hLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50
 5706 CTCTAACAAATTAACCTCCAGGACAGTAGTAACCCCTTATAGACAGCTCG 5755
 51 AsnProHisArgProLeuSerProThrTrpLeuIleIleAspProAspThr 67
 5756 AACCCCATTAACCTTATCCCTTACCTGCGTGAATATGACCTGATAC 5805
 67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84
 5806 GGGGTGCTACGTAAATAGCACTCGAGGTGGTGCCTTAGAGGACCGCTG 5855
 84 rProGlnLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
 5856 GGCTGAACTGCAATTTCTGCTCCGATTTGATTAACCCCGCTGTTAAAGC 5905
 101 ThrProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyThr 117
 5906 ACACCTCCCAACCTAGTCCGATGATTAGGTTATGTGCTGACGACAC 5955
 117 rGluLysGluLysTyrCysGlyGlySerGlyGluSerPheCysArgArg 134
 5956 AGAGAAAGAGAAATAGTGGGGTCTCGGGGAATCCTTCGTAGAGAGAT 6005
 134 rPserCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150
 6006 GGAGCTGCTACCTCCACAGATGAGACGTGGAATGGCCATCTCTCTC 6055
 151 GluAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrLy 167
 6056 CAGACGCGGTAAATTCCTTGTCAATTCGCGCCGCGCAAGTACAA 6105
 167 sMetMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAsp 184
 6106 AATGATGAACATATATAAGATAAGAGCTGCTCCCATCAGACTTAGATT 6155
 184 rLeuLysIleSerPheThrGluLysGlyLysGlnGlnLysAsnIleGlnLy 200
 6156 ATCTAAATATAGTTTCACT...GAAGGAAACACGAAATATTCAAA 6202
 200 sTrpIleAsnGlyMetSerTrpGlyIleValPheTyrLysTyrGlyGly 217
 6203 GTGGATAAATAGTATGAGCTGGGAAATGTTTATATTA..TATGCGCGGG 6251
 217 lValGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGlu 233
 6252 GAGCAGGGTCCACTTAACCATTCGCTTAGGATAGAGACGGGAGACAA 6301
 234 ProProValAlaMetGlyProAspLysValLeuAlaGlnGlnGlyPro 250
 6302 CCCCTGTGGCAATGGAGCCGATAAAGTACTGCTGTAACGAGGGGCCCC 6351
 250 oAlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeu 267


```

|||||
6352 GGGCCGGAGCCGACATTAACCTTGGCCGTCACCAATTAACCTGCTGC 6401
267 rGProaspIleThrgInProProSerAsnSerThrThgIleuIlePro 283
|||||
6402 GGCTGACATTAACAGCGCGCTAGCAACAGTACACAGTGGATTGCTCT 6451
284 ThrAspIleProArgAsnSerProGlyValProValLysThrGlyLna 300
|||||
6452 ACCCAACAGCCCTAGAAACTCCAGAGTGTCTCTTAAAGACAGGACAG 6501
300 gReuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAsp 317
|||||
6502 ACTCTTCACTTCATCCAGGAGCTTCCAGCCATCACTCCACCGAC 6551
317 rGAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTyr 333
6552 CTGATGCCACTCTCTTGTGTGGCTTGTCTATCTCTCAAGGGCTCTTAT 6601
334 TyrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisAr 350
|||||
6602 TATGAGGGGATGGCTAAAGAAATTCATATGACCAAGACATAG 6651
350 gAsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValS 367
6652 AATCAATGTACATGGGGGTCCCGAAATAGCTTACCTCACTGAGATT 6701
367 eRgLYLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeu 383
6702 CCGGAGAGGGGACATGATAGGAAAGCTCCCGCCACCAACACCTT 6751
384 CysTyrSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeuVa 400
6752 TGCTATGACTGTGTATGATGACAGCGCTCAGAAATCAGTATTAGT 6801
400 lProGlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysV 417
6802 ACCTGTTATTAACAGGTGGTGGCATGCAATACGTGGTTAACCCCTGTG 6851
417 alSerThrSerValPheAsnGlnSerLysAspLeuCysValMetValGln 433
6852 TTTCCACCTCAGTCTTCAACCAATCCAAAGATTCTGTGTATGTGCTCA 6901
434 lIleValProArgValTyrTyrHisProGluGluValValLeuAspGluTyr 450
6902 ATCGTCCCGCAGTGTACTACCATCTGAGGAAAGTGGCTGTATGATA 6951
450 rAspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeu 467
6952 TGACATATGGTATTAACGACCAAAAGAAACCCGTATCCCTTACCCCTAG 7001
467 lValMetLeuGlyLeuGlyThrAlaValAlaGlyValGlyThrGlyThrAla 483
7002 CTGTATATCTCGGATTAGGAGCGCGCTGGCGTATGAGACAGGACAGCT 7051
484 AlAlaLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuGluLys 500
7052 GCCCTGATCACAGGACGACAGCAGCTAGGAAAGAGCTGTGTGAGCTACA 7101
500 sAlAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnL 517
7102 TGGCGCCATGACAGAGATCTCCGAGCCCTTAAAGAGCTGTGTGAGCAAC 7151
517 euGluGluSerLeuThrSerLeuSerGluValValIleGlnAsnArgArg 533
7152 TAGAAGAGATCCCTGACTTCTTGTCTGAGAGGTCTTACAGAACCGAGAG 7201
534 GlyLeuAspLeuLeuPheLeuArgGlyGlyLeuCysAlaAlaLeuLys 550
7202 GGATTAGATCTGTGTTCTTAAGAGAAAGGTGGTATGTGACGACCTTAA 7251
550 sGluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSer 567
|||||

```

```

7252 AGAAGAAATGTTGCTTATGATCACTACGAGACCATAGAGACTCCA 7301
567 eTaSnLYsLeuArgLYsLeuGluArgArgArgArgGluVala 583
|||||
7302 TGAACCAAGCTTAGAAAAAGTTAGAGAGGGGTGCAAGAGAAAGAGGCT 7351
584 AspGlnGlyTrpPheGluGlyTyrPheAsnArgSerProTrpMetThr 600
7352 GACCAGGGGTGGTTGAAGATGGTTCACAGAGCTCTCTGGATGACAC 7401
600 rLeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeu 617
7402 CTTCTTTTGTCTGTGAGGGGCCCTAGATGCTCTCTCTCTTACTTA 7451
617 hrValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArg 633
7452 CAGTTGGGCTTGTGTATATAGTTTGTTCCTTGTATAGAAAGCA 7501
634 ValSerAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeu 650
7502 GTGAGTCCAGTCCAGATCATGTAAGGACAGACAGTACCAAGGCTTCT 7551
650 uSerGlnGlyGluThrAspLeu 657
7552 GAGCCAAAGAGAAACTGACCTC 7573
seq_name: gb_v1:AF014162

```

```

seq_documentation_block: 1971 bp mRNA VRL 01-JUN-1998
LOCUS AF014162
DEFINITION Porcine endogenous retrovirus envelope protein (env) mRNA, complete cds.
ACCESSION AF014162
VERSION AF014162.1 GI:3169735
KEYWORDS
SOURCE
ORGANISM
porcine endogenous retrovirus.
porcine endogenous retrovirus
Viruses; Retroid viruses; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
REFERENCE
1 (bases 1 to 1971)
Haworth, C., Galbraith, D.N., Lees, G.M. and Smith, K.T.
Unpublished
JOURNAL
2 (bases 1 to 1971)
Haworth, C., Galbraith, D.N., Lees, G.M. and Smith, K.T.
Direct Submission
TITLE
Submitted (16-JUL-1997) Q-One Biotech Ltd., Todd Campus, Acre Rd.,
JOURNAL
Glasgow G20 0XA, UK
FEATURES
source
location/Qualifiers
1..1971
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
/cell_line="derived from cell line PK15"
/note="PoEV"
1..1971
/gene="env"
1..1971
/gene="env"
/codon_start=1
/product="envelope protein"
/protein_id="AAC1948.1"
/db_xref="GI:3169736"
/translaton="MHPILSRRLPTRGSPKRLRLPISFASIAMFLTLTPQASSK
RLIDSSNPHRLSLTWLILDPDGVYVNSRGVAPRTWMPDELHFCRLRLNPAVKSFP
NLVRSYGFRCPPETREKYGCGSGESFRCRSCVYSNDGDKMPISLDKRVFSFVN
SGPKRYMKLILKXKSCSPSDLYLKTISFTRKTKYSVDKMYELGNSFLKGGGAG
STLIRLRIETGETEPVAMGPDVLAOGPPALPEPNIPLVPLQTLSPDITPPS
TTGLIPNTPRNSPGVYKGTGRLFLIOGAPFALINSTDPATSSCMLCSGPPYE
GMAKERFNTKEHRCNCTWGSNRKLLITFVSGKTCIGAPSHOHLCSYTVYEOA
SENOXLYPGNRMMACNTGLTPCVSYSVFQNSDPCVMYDITVRVYTHHPEVLDLDE
YRINRKRREPVSLAYMLGLTRAVGVTGTLALITGPOOLEKGLGELHAAMEDLKA
LKESVNSLESLTSLSEVVLQNRGLDLFLRGGLCAALKECCFVYDHSGLIRSM
NKLKKLERRRRERADOGWFEWGFNFSPMWTLTLLSATNGLPLVLLTLTVGCLINR

```

BASE COUNT 530 a 490 c 482 g 469 t
 FVAFRRERSAVQIMVLROQYOGILSOGENDL"
 ORIGIN

alignment_scores:

Quality: 3444.50 Length: 658
 Ratio: 5.275 Gaps: 2
 Percent Similarity: 99.240 Percent Identity: 98.936

alignment_block:

US-09-171-553b-10 x AF014162 ..

Align seg 1/1 to: AF014162 from: 1 to: 1971

```

1 MethisProThrLeuSerThrPargHisLeuProThrArgGlyGlyLeuP 17
1 ATGCATCCCAAGCTTAAGCCGGCCACCTCCCGACCTGGGGTGGAGAGCC 50
17 olysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeu 34
51 GAAAGAGCTAGAGATCCCTTAAGCTTGCCCTCCATCGCTGGTTCCTTA 100
34 hLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50
101 CTCTAACATTAACCTCCCGAGCCAGTAGTAAAGCCTTATAGACAGCTCG 150
51 AsnProHisArgProLeuSerProThrIleLeuIleAspProAspThr 67
151 AACCCCAATAGACCTTTATCCCTTACCTGGCTGATTTATGACCTGTATAC 200
67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84
201 GGGTGTCACTGTAATAATAGCACTCGAGGGTGTCTCTCTAGAGCACCCTGGT 250
84 rProGlnLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
251 GGCGTGAACGTGATTCGTGCTCCGATGATTAAACCCGCTGTAAAGC 300
101 ThrProAsnLeuValArgSerThrGlyPheThrCysCysProGlyThr 117
301 ACACCTCCCAACCTAGCTCGGTATGAGGTTCATTCGTCGCCCGACGAC 350
117 rGluLysGluLysTyrCysGlyLysSerGlyLysSerPheCysArgArgT 134
351 AGAGAAAGAGAAATACTGTGGGGGTCTGGGGAATCCCTCTAGAGAGAT 400
134 rPserCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150
401 GGAGCTGCCTCACCTCCCAAGATGAGACTGGAAATGGCCGATCTCTCTC 450
151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyr 167
451 CAGGACCGGGTAATAATTCCTCTTGTCAATTCGGGCCCGGCAAGTACAA 500
167 sMetMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAspT 184
501 AATGATGAAGACTATATAAGATAGAGCTGCCCATCAGACTTAAGATT 550
184 yLeuLysIleSerPheThrGlyLysGlyLysGlnGluAsnIleGlyLys 200
551 ATCTAAGATTAAGTTTCACT...GAAAGGAAACAGGAAATATATCAAAA 597
200 sTrpIleAsnGlyMetSerTrpGlyIleValPheTyrLysTyrGlyLys 217
598 GTGGATTAATGGTATGAGCTGGGGAAATAGCTTTTATTA. TATGGCGGGG 646
217 yAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyLysGlu 233
647 GAGCAGAGCTGCACCTTAACCATTCGCTTAGATAGAGAGCGGAGACAGA 696
234 ProProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProPr 250

```

```

697 CCCCCTGTGGCAATGGACCCGATAAAGTACTGGCTGACAGCGGCCCC 746
250 oAlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuA 267
747 GGCCCTGGAGCCACCGCATTAACCTTCCGGTCCCAATTAACCTTCCGTC 796
267 rProAspIleThrGlnProProSerAsnSerThrThrGlyLeuIlePro 283
797 GGCCGACATACACACAGCGGCTAGCAACATACACTGAGTATATTCCT 846
284 ThrAsnThrProArgAsnSerProGlyValProValLysThrGlyLys 300
847 ACCAACAGCGCTAGAAACTCCCGAGGTGTTCTGTAAAGACAGACAGAG 896
300 gLeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAsp 317
897 ACTCTTCAGTCTCATCCAGGAGGCTTTCAGGATCAATCAATCCACGACC 946
317 rAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTyr 333
947 CTGATGCCACTTCTTCTTGTGGCTTGTATCTCATCTCAGGGCCCTTAT 996
334 TyrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHis 350
997 TATGAGGGGATGGCTAAAGAAAGAAATTCATGTGACCAAGAGCATAG 1046
350 gAsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluVal 367
1047 AAATCAATGTACATGGGGGTCCGAAATTAAGCTTAACCTCAGCTAACTTT 1096
367 erGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeu 383
1097 CCGGGAAGGGGACATGATAGAAAGAACTCCCACTCCACACACACTT 1146
384 CysTyrSerThrValValTyrGluGlnAlaSerLysGlnIleTyrLeu 400
1147 TGTATATAGTACTGTGTTATAGACAGCGCTCAGAAATACATGATTTAGT 1196
400 lProGlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCys 417
1197 ACCGTGTTATACAGGTGGTGGCATGATCTAGGGTTAAACCCCTGTG 1246
417 aLserThrSerValPheAsnGlnSerLysAspLeuCysValMetValGln 433
1247 TTTCACACTGACTCTTCACCAATCCAAAGATTCTGTGCATGATGCCAA 1296
434 lLeuValProArgValTyrTyrHisProGluGluValValLeuAspGlu 450
1297 ATCGTCCCGCAGTGTACTACCATCTGAGGAAGTGTCTGTATGATA 1346
450 rAspTyrArgTrpAsnArgProLysArgGluProValSerLeuThrLeu 467
1347 TGACTATCGTATACAGGACCAAAAGAGAACCCGTTATCCCTTATACCT 1396
467 lValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAla 483
1397 CTGTAATAGCTGATAGGAGCGCGTGGCGTAGAGAACAGGACAGCT 1446
484 AlaleuIleThrGlyProGlnGlnLeuGluLysGlyLeuGlyLeuHis 500
1447 GCCCTGATCACAGGACACAGCAGCTAGAGAAAGAGCTGTGTGAGCTACA 1496
500 sAlaAlaMetThrGlnAspLeuArgAlaLeuLysGlySerValSerAsn 517
1497 TGGCGGCATGACAGAGATCTCCGAGCCTTAAAGAGATCTGTATGCAACC 1546
517 euGlnGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 533
1547 TAGAAGAGTCCCTGATCTTCTTGTGAAGGTGTTCACAGAACCGGAGG 1596
534 GlyLeuAspLeuLeuPheLeuArgGluGlyLysCysAlaAlaLeuLys 550
1597 GGATTAGATCTGCTGTTCTAAGAGAGGTGGTTATGTGACAGCTTAAA 1646

```

550 sglunglucyscyspheytyvalasphhisserglyalalilearagspserm 567
 1647 ACAAAAGATGCTTCTTATGATGATCAAGAGACCACTGAGACACTCA 1696
 567 etasnlylsleuargylsleuargyluarargyluaragluala 583
 1697 TGAACACACTTGAAGAAAGTAGAGAGCGCTCGAAGGAAAGAGAGAGCT 1746
 584 aspginglytrpheygluglytyrphesnasrgserprotpmetthth 600
 1747 GACCGAGGGGTGGTTGAAGAGATGTTCAACAGGCTCTCTTGATGACAC 1796
 600 rleuenseraleutrhnglyproleuvalleuenserleuenserleut 617
 1797 CCGTCTTCTGCTGAGAGGCGCCCTGATGATCTCTCTCTGATCTTA 1846
 617 hrvalglyprocysleuileasnarqphvalalaphvalargyluar 633
 1847 CAGTTGGGCGCTTGAATTAAGTGTGTTGCTTGTAGAGAACGA 1896
 634 valseralavalglnilemetvalleuarlgnglntyrngllyleule 650
 1897 GTGAGTGCAGTCCAGATCATGCTACTAGGCAACAGTCAAGGCTTCT 1946
 650 userglnglyluthrleu 657
 1947 GAGCCAGAGAGAACTGACCTC 1968

seq_name: gb_pat:A66553

seq_documentation_block:

LOCUS A66553 8209 bp DNA PAT 29-MAR-1999
 DEFINITION Sequence 3 from Patent WO9740167.
 ACCESSION A66553
 VERSION A66553.1 GI:4538106
 KEYWORDS
 SOURCE porcine endogenous retrovirus.
 ORGANISM porcine endogenous retrovirus
 Viruses: Retroviridae; Mammalian type C
 retroviruses; 1-Mammalian type C virus group.

REFERENCE
 1 (bases 1 to 8209)
 Galbraith,D.N., Haworth,C., Lees,G.M. and Smith,K.T.
 TITLE PORCINE RETROVIRUS
 JOURNAL Patent: WO 9740167-A 3 30-OCT-1997;
 Q ONE BIOTECH LTD (GB)

FEATURES
 Source
 1..8209
 Location/Qualifiers

CDS
 /organism="porcine endogenous retrovirus"
 /db_xref="taxon:61673"
 588..2162
 /product="1"
 /protein_id="CAB39347.1"
 /translation="MGQTVTPLSLTLDHMTVEYRAHNLISVQVKKGMQTCASEMP
 TFDVGPSESTENSEIILAKIIFQTPGSHDQEPYLLIMODLADPPVWKPMNLN
 KPRKPGRIALGEKNHSAKEVPESSYLPDRGADLAGTPTSCPTLSGCEG
 TSAPGAPVVEGACGRSRGATPERTDIALPLRTGPPMPGGLOLPLOMPSS
 ADLYNKNTHNPSPEDPORLTGLVESIMFSHOHTDCCOOLLOTLETERREILIEA
 RKVPGADGPTOLQNEIDMGFPLTRPGMDYNTARESEKIKYRIALVAGLSASRP
 TNLAKREVWQGNPEPVSFLERLMEARFRPTEDTSEOKASVALAFIGQSALDIR
 KKLQLEGLQEAELRDLYRAEKVYRRETEBEKEDRKEREERERDRQEKMLT
 KTLAAVEGKSSREREDFRKIRSGPROSGNLGNRTPLDKQCAVCYCKEKGHMARNCPK
 KGNKGPVALLEEDKXD"

CDS

<2163..5747
 /codon_start=1
 /product="REVERSE TRANSCRIPTASE"
 /protein_id="CAB39348.1"
 /db_xref="GI:4538108"
 /translation="GRRGSDPLPEPRVTLKVEGQVPEPLVDGAEHSVLLQPLGKLE
 KSKMVGATGQROYPTWTRRTYDLGVRGRTYHSLVYPECPYPLGDLTKMAQSLKE
 EQGRPEVSNNKPTIVTLTLDDEYRLYSPOVAPDDIDIOSWLEQFOVAAETAGKGLA

KQVPOVYIOLKASATPVSVROYPLSREAREGIMPHVQRLIQOGLIVPQSPWMTPLP
 VRKPTNDYRVPQDLREYNKRVQDILHPVPNPYMLSLPPERMWYVLLDKAFEL
 RLKPSOPLFAFEMRDPCTGRTGQULWTRLPQGFKNSTPTEDEALHDLNFRLOHQR
 VTILOYVDILLACATKODCEGKALLLEISDYGASAKAQICREYTYLGSYLR
 GGORLUTPARKKTYVQUPAPPTAKQVRELESTAFCSYLMIPGATLAPYLTKEG
 GFSWPEHOKAFDITKALSLAPALBDEVTKPTTLVDEKQARKVLOTLPNR
 PVAYLSKRLDVASGWPCYKALIAVALILVNDADKLLGQVITVIAHALENTIROP
 DRMMNARMATYOSLLITERTPAPALNPVTLPEBETEPVPHDOLLIEBTGVR
 KDLDIPLTGEVLTWTFDSSVVEGKMACAAVADGRTIMASLPEGTSAQAEIL
 ALTOALRIAEKSNINYDSSYAFATAHVGAIVKORGLTSAGREIKNKEILSILE
 ALHLPKRLAITHCGHOKAKDLISRNOMDRVAKQAQAVNLLPIETPKAPPRRO
 YLLEWDEITKIDPSETBCTYSTGKELPKEBDEIVOOITRHLITCKHLQOL
 VRTSPYHVLRLPGVADSVYKVCQOLNANPSRIPEGRRLGSHPGAHWEVDTEVK
 PAKYGNKYLAVLVTFSGWVAYPTKETSVAKKLIEELFPFRGIPKVIQSDNGPA
 FVAQYOSGLAKILGIDMKLHCAYRPOSSGOYERNNRTIKETLITLTTETGINDMALL
 PVLFRVNTPTGOGILPEYELVLTGGPPLAIFARASADYVLSQPLSRKALEWVR
 RAMQOLRAIYSGSDLPVHRQYVDSVYRRHAGNLETRKPGYLVLLITTPRAVKE
 GIPVTHASHVYPAPPDPSGWRKATEMPLKRLHLRPLVSNNSPGO"

CDS

/codon_start=1
 /product="VIRAL ENVELOPE PROTEIN"
 /protein_id="CAB39349.1"
 /db_xref="GI:4538109"

/translation="MHPILSRHLPTNGEPRKRLRIPLSASTAMPLITLITTPQASK
 RLIDSSNPHRLSLTWLIIIDPTGVTVNASTGVARPTWPELHFCRLINPAKSTP
 PNLVRSYFYCCPGTEKEXYCGSGSECTVSCVNDGDKWMPISLQIRVKSFPN
 SGPGYKMKMLYKDKSCSPDLXLISFTERRKTKSKYDKVWELNSPFLYGGAG
 STTIRLRIETGTPPVAMGPDKVLAEGPPALPBPILNPOLTSRLDITOPSS
 TTGLIPMTNPNSPGVYKTKGQRLPESLQGAFOAINSTDPDAPSCMLCSGPPYE
 GMAKERRKNTVKEHNOCTWGSRNKLTILEVSGGTCTIGARPSHOILCTSYVEA
 SENDVLPGVIRNMWACNTGLTPCVSTSVFNOSKFCVAVQVLPVRYVHPEVVLDEID
 YRYNRPKREPVSLTAVMLGIGTAVGVTGIALITLPQOLEKGLAHAMTEDLA
 LKESYNLEESLTLSSEYVLQNRGLDILFIREGLCALKEBCCFYVDHSGATRDSM
 NLRKLELRERRERADQGEWGFENRSPMTTILSALITGLVLVLLILLITVGPCLINR
 FVAFYREVSVAVOIMVLNQYQOGLISGETDL"

BASE COUNT
 ORIGIN
 2168 a 2064 c 2154 g 1823 t

alignment_scores:
 Quality: 3444.50 Length: 658
 Ratio: 5.275 Gaps: 2
 Percent Similarity: 99.240 Percent Identity: 98.936

alignment_block:
 US-09-171-553B-10 x A66553 ..
 Align seg 1/1 to: A66553 from: 1 to: 8209

1 MethisProthrleuSerTrpArgHisleuProthrArgglylyglubr 17
 5620 ATGCATCCACGTTAAAGCCGCGCCACCTCCGACCTGGGGTGAAGACC 5669
 17 olyArgleuArggleProleuSerPheAlaSerIleAlaTrpPheLeu 34
 5670 GAAAGACTGAGATCCCTTAAGCTTGGCTTCATCGCTTGGTCTTA 5719
 34 hrleuthrilethrProglAlaSerSerlysrArgleuileasPserSer 50
 5720 CTCGTAACAATACATCCCGACGAGGACAGTAGTAACGCTTATAGACACCTCG 5769
 51 AsnProHisarProleuSerProthrTrpPheulleleasPProaspth 67
 5770 AACCCCATAGACCTTATCCCTTACCTCGCTCATATTATGACCCCTGATAC 5819
 67 rglyValThrValasnerThrArgglyValAlaProArglyThrTrp 84
 5820 GGGGTACTGTAAATAGACCTGAGGTGTGTCTCTAGAGGACCTGGT 5869
 84 rpprogluleuHisPheCysleuArgleuileasnarProAlaValysSer 100
 5870 GGCCTGACATGCAATTTCTGCTCGCATTTGATTAACCCCGCTGTAAAGC 5919

```

101 ThrProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyIth 117
|||||
5920 ACACCTCCCAACCTAGTCCGATGATGGCTTCTATGCTGCCAGGCAC 5969
|||||
117 TGIuysgluysTyrCysGlyIysSerGlyGluSerPheCysArgArgT 134
|||||
5970 AGAGAAAGAAAGAAATCTGGGGGGTCTGGGGAATCCTCTGTAGAGAT 6019
|||||
134 rPserCysValThrSerAsnAspGlyAspTlpLysTTPProIleSerLeu 150
|||||
6020 GGAGCTGGCTACCTCCAAACGATGGAGACTGGAAATGGCCGATCTCTCTC 6069
|||||
151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrLy 167
|||||
6070 CAGGACCGGGTAAATCTCTTGTCAATTCGGCCCGGCGCACTACAA 6119
|||||
167 smetLeuLysLeuTyrLysAspLysSerCysSerProSerAspLeuAspT 184
|||||
6120 AATGATGAACATATATAAGATTAAGAGCTGCCATCCCATCTAGATT 6169
|||||
184 yrlEulysIleSerPheThrGlyLys.GlyLysGlnGluAsnIleGlnLy 200
|||||
6170 ATCTAAAGATAAGTTTCACT...GAAGGAAACAGGAAATATTCAGAA 6216
|||||
200 strPileasnglymetSerTTPGlyIleValPheTyrLysTyrGlyGly 217
|||||
6217 GTGGATTAATGCTATGAGCTGGGGAATAGTTTATATA.TATGGCGGG 6265
|||||
217 lYalaglySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGlu 233
|||||
6266 GAGCAGGGGTCCACTTAAACCATTCGCTTAGATAGACGGGGGACAGAA 6315
|||||
234 ProProValAlaMetGlyProAspLysValLeuAlaGlnGlnIleProPr 250
|||||
6316 CCCCCTGGTGCAATGGAGCCGATTAAGTACTGGCTGAACAGGGGCCCC 6365
|||||
250 oAlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuA 267
|||||
6366 GGCCCTGGAGGCCACCGCATACTTGGCGTGGCCCATTAACCTGGCTGC 6415
|||||
267 rPProAspIleThrGlnProProSerAsnSerThrThrGlyLeuIlePro 283
|||||
6416 GGCGTACATTAACACAGCCGCTACCAACAGTACACATGATGTATGCT 6465
|||||
284 ThrAsnThrProArgAsnSerProGlyValProValLysThrGlyGlnAr 300
|||||
6466 ACCCAACGCCCTAGAAACTCCCAAGTCTTCTGTTAAGACAGGACAGAG 6515
|||||
300 gLeuPheSerLeuIleGlnIleAlaPheGlnAlaIleAsnSerThrAspP 317
|||||
6516 ACTCTTCATCTCATCCAGGAGCTTCCAAACCATCAACTCCAGCCGACC 6565
|||||
317 roAspAlaThrSerSerCysTTPLeuCysLeuSerSerGlyProProGlyr 333
|||||
6566 CTGATGCCACTCTCTTGTGGCTTGTCTATCTCTCAAGGCCCTCTAT 6615
|||||
334 TyrGlnGlyMetAlaLysGlnArgLysPheAsnValThrLysGlnHisAr 350
|||||
6616 TATGAGGGGATGGCTTAAGAAAGAAATTCATATGACCAAGAGCATAG 6665
|||||
350 gAsnGlnCysThrTTPGlySerArgAsnLysLeuThrLeuThrGluValS 367
|||||
6666 AATTCATATGATGGGGGCCGGAATTAAGCTTACCTCACTCAAGATT 6715
|||||
367 erGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeu 383
|||||
6716 CCGGGAAGGGGACATGATGAAAGCTCCCCATCCCAACACCTT 6765
|||||
384 CysTyrSerThrValValTyrGlnGlnAlaSerGlnAsnGlnTyrLeuVa 400
|||||
6766 TGCATATAGTACTGTGTTATATGACGAGGCTCAGAAATACATATTATAT 6815
|||||
400 lProGlyTyrAsnArgTTPTrpAlaCysAsnThrGlyLeuThrProCysV 417
|||||
8616 ACCGTGTTATACAGGTGTCGCGCATGCAATATCTGGGTTAACCCCTGTG 8665
|||||
417 alSerThrSerValPheAsnGlnSerLysAspLeuysValMetValGln 433
|||||
6866 TTTCCACCTAGTCTTCAACCAATCCAAAGATTTCTGTGTCATGTCGAA 6915
|||||
434 lLeValProArgValTyrTyrHisProGlnGluValValLeuAspLufTy 450
|||||
6916 ATGCTCCCCGAGTGTACTACATCCCTGAGCAAGTGGTCTCTATGATAA 6965
|||||
450 rAspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuA 467
|||||
6966 TGACTATCGGTATTAACCGACCAAAAGAGAACCCGTATCCCTTAACCTAG 7015
|||||
467 lAlaMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAla 483
|||||
7016 CTGTAAATGCTCGGATTAAGGAGCGCGTGGCGTAGAGAACAGGACAGCT 7065
|||||
484 AlaleuThrThrGlyProGlnGlnLeuGlnLysGlyLeuGlyLysLeuH 500
|||||
7066 GCCCTGATCACAGGACCAACAGAGCTAGAGAAAGAGCTGTGAGCTACA 7115
|||||
500 salAlaMetThrGlnAspLeuArgAlaLeuLysGlnSerValSerAsnL 517
|||||
7116 TGCGCCATGACAGAAAGATCTCCAGCCTTAAAGAGACTCTGTAGCAACC 7165
|||||
517 euGlnGlnSerLeuThrSerLeuSerGlnValValLeuGlnAsnArgArg 533
|||||
7166 TAGAAGAGTCCCTGACTTCTTGTCTGAAGTGTGTCTACGAAACCGAGG 7215
|||||
534 GlnLeuAspLeuLeuPheLeuArgGlyGlyLysCysAlaAlaLeuLy 550
|||||
7216 GGATTACATCTGCTGCTTCTTAAGAGAGTGGCTATGTGACAGCTTAA 7265
|||||
550 sGlnGlnCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerM 567
|||||
7266 AGAAGAAATGTTGCTTATGTAGTACACTCAGGAGCCATCAGAGACTCA 7315
|||||
567 eLAsnLysLeuArgLysLysLeuGlnArgArgArgGluArgGluAla 583
|||||
7316 TGAACAAAGCTTAGAAAAAGTTAGAGAGCGCTCGAAGGAAAGAGAGCT 7365
|||||
584 AspGlnGlyTTPPheGlnGlyTTPPheAsnArgSerProTTPmetThrTh 600
|||||
7366 GACCAGGGGTGTTTGAAGATGGTTCACACAGCTCTCTGATGAGACCAC 7415
|||||
600 rLeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuT 617
|||||
7416 CCTGCTTCTGCTCTGACGGGGCCCTAGTAGTCTCTCTCTCTGTTACTTA 7465
|||||
617 hrValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArg 633
|||||
7466 CAGTTGGGCTTGTCTTAATTAATAGTTTGTGGCTTGTGTAGGAAGCA 7515
|||||
634 ValSerAlaValGlnIleMetValLeuArgGlnIleTyrGlnGlyLeu 650
|||||
7516 GTGAGTGCAGTCCAGATCATGTACTTAGCCAAAGTACCAAGGCTTCT 7565
|||||
650 uSerGlnGlyLysThrAspLeu 657
|||||
7566 GAGCCAAAGAGAACTGACCTC 7587
|||||
seq_name: gb_v1:PEN28589
seq_documentation block:
LOCUS PEN28589 1968 bp RNA VRL 29-AUG-2000
DEFINITION Porcine endogenous retrovirus PERV-B1 env gene for envelope
protein, genomic RNA.
ACCESSION AJ28589
VERSION AJ28589.1 GI:7414447
KEYWORDS env gene; envelope protein.
SOURCE porcine endogenous retrovirus.

```

ORGANISM porcine endogenous retrovirus
 Viruses; Retrovirdae; Mammalian type C
 retroviruses; 1-Mammalian type C virus group.
 REFERENCE 1 (bases 1 to 1968)
 AUTHORS Bosch, S., Arnaud, C. and Jestin, A.
 TITLE Study of full-length porcine endogenous retrovirus genomes with
 envelope gene polymorphism in a specific pathogen-free large white
 swine herd
 JOURNAL J. Virol. 74 (18), 8575-8581 (2000)
 MEDLINE 20411432
 REFERENCE 2 (bases 1 to 1968)
 AUTHORS Bosch, S.
 TITLE Direct Submission
 JOURNAL Submitted (28-MAR-2000) Bosch S., Molecular Biology Department,
 AFSSA-zoopole, Zoopole Les Croix B.P.53, Ploufragan 22440, FRANCE
 FEATURES
 source
 1..1968
 /organism="porcine endogenous retrovirus"
 /specific_host="Sus scrofa"
 /db_xref="taxon:61673"
 /clone="PERV-B1"
 /note="French Large White SPF-pig"
 1..1968
 /gene="env"
 1..1968
 /gene="env"
 /codon_start=1
 /product="envelope protein"
 /protein_id="CAB86208.1"
 /db_xref="GI:7414448"
 /translation="MPTLSMRQPTRGCEPRKRLIPLSFSGIAMEFLITTPQSSK
 RLIDSNPHRLSLTWLITDPDGVNSTKVAAPGTPWMLHCLRLINAVSTP
 PNLVRSYGYCCGCEKERYCGSGESFCRWSCVISNGDKWPLSLDDRYKSEFN
 SGGRKVMKLYKDKSCSPSDLYLKISFTEKGKNIKMMGMSWGLVFKYGGGA
 GSTLTIETGTEPPVAVGDPKVLAEQGPALPPHNLVPOLTSRPDITOPSGMT
 GLIPTIPRSPGVYKTVGSLQGAFOAINSTDPAATSCWMLSSGGPYEGM
 AKGKFNVTKEHNOCTGSRNKLITLTVSGGCTGCGAKRPHQHLCTSYADEASE
 NOYLPGVNRWMACTGTPCVSTSTFENOSKPCVAVQVLPKRYTHPEBMVIDEYR
 YNRKREPVSLLIAVNLGLTAVGVTGAAXITGPQLEKLGELHMAITDLALE
 ESYNLSESLTSLSEVNLNRGLDLFLREGCLALKEBCCFVVDHSGAIRDMSK
 LRELRERRERADQGFEGMFWNRPMTTLLSLALTGVLVLLLTIVGCLINRFV
 AFVREKVSVOIIMVLRMLKGEFOHTGGRY"
 BASE COUNT 522 a 483 c 492 g 468 t 3 others
 ORIGIN
 alignment_scores:
 Quality: 3340.50 Length: 645
 Ratio: 5.252 Gaps: 1
 Percent Similarity: 98.605 Percent Identity: 96.124
 alignment_block:
 US-09-171-553B-10 x PEN288589 ..
 Align seg 1/1 to: PEN288589 from: 1 to: 1968
 1 MethSPROTHrLeuSerTTPaTHHisLeuProTHrArgLylGlyLupr 17
 1 ATGCAATCCACGTTAAGCTGCGCCAAACCCGACCTCGGGGTGGGAGCC 50
 17 oLysArgLeuArgLLeuProLeuSerPheAlaSerLeuAlaLTPheLeu 34
 51 GAAGAAGCTGGGAATCCCTTAACTGGCTCCATCGCTGGTCTCTTA 100
 34 hrLeuThrLeuThrProGlnAlaSerSerLysArgLeuLeuLeuSerSer 50
 101 CTCTNACAAATAACTCCCGAGCCAGTAGTAAGCCCTTATAGACAGTCG 150
 51 AsnProHisArgProLeuSerProTHrTTPLeuLlelleasProAspTh 67
 151 AACCCCATAGACCTTATCCCTTACTGCTGATATTATGACCCCTATAC 200
 67 rGlyValThrValAsnSerThrArgLylValAlaIaProArgGlyThrThr 84

201 GGGGTGCTACTGTAAATAGCACTGGGGGTGTTCCTTCCATAGAGGACCTGGT 250
 84 rPrProGlnLeuHisPheCysLeuArgLeuLeuAsnProAlaValLysSer 100
 251 GGCTTGACATGCAATTCCTGCTCCGATGATTAACCCCGCTGTAAAGCC 300
 101 ThrProProAsnLeuValArgSerThrGlyPheThyCysCysProGlyThr 117
 301 ACACCTCCCAACCTAGTCCTGATGTTATGGTTCTATTGCTGCCAGGAC 350
 117 rGlyLysGluLysTyCysGlyLysSerGlyLysLeuPheCysArgArg 134
 351 AGAAGAAAGAAATACTGTGGGGTTCGGGGAATCCTTCTGTAAGAGAT 400
 134 rPseSerValThrSerAsnAspGlyAspThrLysThrProIleSerLeu 150
 401 GGAGCTCGTCCACCTCCACAGATGAGACGGAATGGCCGATCTCTTC 450
 151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysThr 167
 451 CAGAGCCGGGTAAATTCCTCTTGTCAATTCGGGCCGGGCAAGTACAA 500
 167 sMetMetLysLeuThrLysAspLysSerCysSerProSerAspLeuAsp 184
 501 AGTGATGAACCTATATTAAGATTAAGAGCTGCTCCCATCAGACTTAGATT 550
 184 rYrLeuLysIleSerPheThrGluLysGlyLysGlnLysIleGlnLys 200
 551 ATCTAAAGATTAAGTTTCACGAAAGAAAGAAAGAAATATATCAAAAG 600
 201 TrpIleAsnGlyMetSerTrpGlyIleValPheThyLysTyGlyLys 217
 601 TGGATGATGTATGAGCTGGGGAATAGTTTTTATTAATATAGCCGGGG 650
 217 yAlaGlySerThrLeuThrIleArgLeuArgLleGlyThrGlyLup 234
 651 AGAGAGGTCCACTTAACCAAT.....ATAGAGAGCGGGGACAGAAC 691
 234 rPrProValAlaMetGlyProAspLysValLeuAlaGlnGlyProp 250
 692 CCCCTGTGGAGTGGAGACCCGATTAAGTCTGCTTAAACAGGAGGCCCGG 741
 251 AlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeu 267
 742 GCCCTGAGACCGCATTAAGTTCGGGTGCCCATTAACCTCGGTGG 791
 267 gPrAspIleThrGlnProProSerAsnSerThrThrGlyLeuIlePro 284
 792 GCCTGACATTAACACACGCCCTTACGACGCTACCATGATTCGATTCCTA 841
 284 hrAsnThrProArgAsnSerProGlyValProValLysThrGlnArg 300
 842 CCAACACGCTTAGACACTCCCGAGGTGTTCTGTTAAGCAGACAGACAGA 891
 301 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAsp 317
 892 CTTCTAGCTCATCCAGGAGCTTCCAAAGCATCACTCCAGCAGACC 941
 317 oAspAlaThrSerSerCysTrpLeuLysLeuSerSerGlyPropoty 334
 942 TATGCGCACTCTTCTTGTGCTTGTCTATCTACCTAGGCGCTCTCTATT 991
 334 rGlyGlyMetAlaLysGluArgLysPheAsnValThrLysGlnHisArg 350
 992 ATGAGGGGATGCTTAAGAGAGAAATTCATAGTACCAAGAGCATGAGA 1041
 351 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuGlyValSe 367
 1042 AATCATATGACGTGGGGGTCGGAATTAAGCTTACCCTCAGTAAGTTTC 1091
 367 rGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeu 384

```

VERSION      AJ288588.1 GI:7414445
KEYWORDS     env gene; envelope protein.
SOURCE       porcine endogenous retrovirus.
ORGANISM     porcine endogenous retrovirus
REFERENCE    1 (bases 1 to 2000)
AUTHORS      Bosch, S., Arnaud, C. and Jestin, A.
TITLE        Study of full-length porcine endogenous retrovirus genomes with
             envelope gene polymorphism in a specific pathogen-free large white
             swine herd
JOURNAL      J. Virol. 74 (18), 8575-8581 (2000)
MEDLINE      20411432
REFERENCE    2 (bases 1 to 2000)
AUTHORS      Bosch, S.
TITLE        Direct Submission
JOURNAL      Submitted (28-MAR-2000) Bosch S., Molecular Biology Department,
             AFSSA-Zooopole, Zoopole Les Croix B.P.53, Ploufragan 22440, FRANCE
FEATURES     Location/Qualifiers
             1..2000
             /organism="porcine endogenous retrovirus"
             /specific_host="Sus scrofa"
             /db_xref="taxon:61673"
             /clone="Perv-AB7"
             /note="Perv Large White SPF-pig"
             1..606
             /gene="env"
             1..606
             /note="truncated, not functional"
             /codon_start=1
             /product="envelope protein"
             /protein_id="CA86207.1"
             /db_xref="GI:7414446"
             /translation="MHPTLSMNHLPTRGGEPRRLRIPSPASIMFLUTLTPPOASSK
             RLDSNPNRRLSTLWLIIDPDGTGVSTGVAARGTWPELHRCRLIPAVKSTP
             PLVLSYGFYCCPGTEKERYCGSGSESCSRKSCVTSMDGIMKPIISLDYVKFSFVN
             SGPGRYKWKLYKDKSCSPSDLDYLFKISFTFKGRKIRKSG"
BASE COUNT   532 a 496 c 499 g 473 t
ORIGIN
alignment_scores:
  Quality: 3340.00      Length: 646
  Ratio: 5.219          Gaps: 1
  Percent Similarity: 99.071  Percent Identity: 96.749
alignment_block:
  US-09-171-553B-10 x PEN288588 ..
Align seg 1/1 to: PEN288588 from: 1 to: 2000
1 MethSPProthLeuSerThrPARGHISLeuProthArgGlyGlyIupr 17
  |||||||
1 ATGCATCCACGTTAAAGCTGGCGCCACCTCCCGACTGGGGGTGAGAGCC 50
  |||||||
17 OlyArgLeuArgGlyLeuProLeuSerPheAlaSerIleAlaIleProPheLeuT 34
  |||||||
51 GAAGAAGGTGAGAAATCCCTTAAGCTTCGCTCCATCGCTGCTGCTTAA 100
  |||||||
34 hLeuThIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50
  |||||||
101 CTCTAACATTAACCTCCAGGCGAGTGAAGTAAACCTTTATAGACAGCTG 150
  |||||||
51 AsnProHISArgProLeuSerProThrTrpLeuIleIleAspProAspTh 67
  |||||||
151 AACCCCATAGACCTTTATCCCTTACCGGCTGATATGACCTGATAC 200
  |||||||
67 rGlyValThrValAsnSerThrArgGlyValAlaIleProArgGlyThrTrrp 84
  |||||||
201 GGGTGTCACTGTAATAATAGACTGAGAGTGTGCTCCTAGAGGACACCTGT 250
  |||||||
84 rProGlnLeuHISpHeCysLeuArgLeuIleAsnProAlaValLysSer 100

```

```

seq_name: gb_v1: PEN288588
seq_documentation_block:
LOCUS       PEN288588      2000 bp      RNA
DEFINITION  Porcine endogenous retrovirus Perv-AB7 env gene for envelope
protein, genomic RNA.
ACCESSION   AJ288588

```

```

251 GGCCTAACTGCAATTTCTGCTCCGATGATTACCCGCTGTAAAC 300
101 ThrProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117
301 ACACTCCCACTAGTCGCTAGTATGAGTCTCTATGCTGCGCAGCAC 350
117 rGluGlySerGlyTyrCysGlyGlySerGlyGlySerPheCysArgArg 134
351 AAGAGAAAGAAATACTGTGGGGTCTGGAGAAATCTCTCTAGAGAT 400
134 rSerCysValThrSerAsnAspGlyAspTrpLysTrpProLysLeu 150
401 GAGAGCTGCTCACTCCCAACGATGAGACTGGAATGGCCGATCTCTC 450
151 GluAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyr 167
451 CAGGACCGGGTAAATTCCTCTTCTCAATTCGCGCGGGAAGTACAA 500
167 sMetLeuLysLeuTyrLysAspLysSerCysSerProSerAspLeuAsp 184
501 AGTGAATGAACCTATATAAAGATGAAGCTGCTCCCATCAGACTTAGAT 550
184 rLeuLysLysLeuSerPheThrGluLysGlyLysGluAsnLysLeu 200
551 ATCTAAGATAGTTTCACTGAAAGAGAA. CAGGAAATATTTCAAG 599
201 TrpIleAsnGlyMetSerTyrPglyLysValPheTyrLysTyrGly 217
600 TGGATTAATGTATGAGCTGGGGAATAGTTTATTAATATGGCGGG 649
217 yAlaGlySerThrLeuThrIleArgLeuArgLysGluThrGlyThrGlu 234
650 AACAGGCTCCACTTAACCATTCGCTTGAAGATGAGACGGGACAGAC 699
234 rProValAlaMetGlyProAspLysValLeuAlaGluGluGlyProPro 250
700 CCCCTGGGCACTGGGAGCCGATTAAGTACTGGCTGAACAGGGCCCCG 749
251 AlaLeuGluProProHisAsnLeuProValProGluLeuThrSerLeu 267
750 GCCCTGGAGCACCGCATTAACCTCCGCTCCCAATTAACCTCGCGCG 799
267 gProAspIleThrGlnProProSerAsnSerThrThrGlyLeuLeuPro 284
800 GCCTGACTTAACACAGCCGCTACCAACGATCAGTGAATGATTCCTA 849
284 hrAsnThrProArgAsnSerProGlyValProValLysThrGlyLys 300
850 CCAACAGGCTTAAAGAACTCCCGAGGTCTCTGTTAAGACAGACAGAGA 899
301 LeuPheSerLeuIleGln.GlyAlaPheGlnAlaIleAsnSerThrAsp 317
900 ATCTTCAGTCTCATCCAGGGAGCTTCCAAAGCCACCACTCCACGAC 949
317 rAspAlaThrSerSerCysTrpLeuGlySerLeuSerGlyProProTyr 333
950 CTGATGCCACTCTCTTGTGTGCTTGTATCTCAGGGGCTCTTAT 999
334 TyrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHis 350
1000 TATGAGGGAGTGGCTAAAGAGAAATTTAAAGTGACCAAGAGCATAG 1049
350 gAsnGluCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluVal 367
1050 AAATCAATGTACATGGGGGTCCCAATAAGTTACCTCAGCAAAATTT 1099
367 erGlyLysGlyThrCysLysGlyLysAlaProProSerHisGlnHisLeu 383
1100 CCGGGAAGGGGACATAGAGAAAGCTCCCATCCACCAACCACTT 1149
384 CysTyrSerThrValValTyrGluGlnAlaSerGluAsnGluTyrLeuVal 400
1150 TGTATAGTACTGTGTTATGACAGGCTCCAGAAATAGATATTAGT 1199
400 lProGlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCys 417
1200 ACGTGGTTGACAGAGGTGGTGGCATCATATACGAGGTAAACCCCTGT 1249
417 alSerThrSerValPheAsnGlnSerLysAspLeuCysValMetValGln 433
1250 TTTCACCTCAGCTTCAACCAATCCAAAGATTTCTGTGATGATGCCA 1299
434 lLeuValProArgValTyrTyrHisProGluGluValValLeuAspLys 450
1300 ATCGTCCCGAGAGTACTACCATCCGAGGAAGTGGTCTTATGATGA 1349
450 rAspTyrArgTyrAsnArgProLysArgGlyLysProValSerLeuThrLeu 467
1350 TGACTATCGGTATACAGGACCAAAAGAACCGATGCCCTTAACCTAG 1399
467 lValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAla 483
1400 CTGTAATGCTCGGATTAGGAGCGCGCTTGGCTAGGAACAGGACAGCT 1449
484 AlaLeuIleThrGlyProGluGlnLeuGluLysGlyLeuGlyLysLeu 500
1450 GCCCTGATCACAGAGACACACAGACTAGAGAAAGACTTGTGACCTACA 1499
500 sAlaAlaMetThrGluAspLeuArgAlaLeuLysGlySerValSerAsn 517
1500 TGGCGTATATACAGAAAGATTCGAGACCTTAGAGAGTCTGTATGACAC 1549
517 euGluGluSerLeuThrSerLeuSerGlyValValLeuGluAsnArgArg 533
1550 TAGAAGAGTCCCTGACTTCTTGTCTGAAGTGGTCTTACAGAAACGGAG 1599
534 GlyLeuAspLeuLeuPheLeuArgGlyGlyGlyLeuCysAlaAlaLeu 550
1600 GGAATGATCTGCTGTTCTTAAAGAAAGTGGGTTATGTCAGCCTTAA 1649
550 sGluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSer 567
1650 AGAAGATGTGCTTCTATGTATGATCCTCAGAGGACCATCAGACTCCA 1699
567 etAsnLysLeuArgLysLysLeuGluArgArgArgLysGluArgLys 583
1700 TGAGCAAGCTTGAAGAAAGTTAGAGAGCGTCGAAGGAAGAGAGGCT 1749
584 AspGluGlyTyrPheGluGlyTyrPheAsnArgSerProTyrMetThr 600
1750 GACCAAGGGTGGTTGAAGGATGGTTCAACAGGTCTCTTGGATGACAC 1799
600 rLeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeu 617
1800 CCGCTTGTGCTCTGACGGGACCCCTAGTACTCTCTCTGTTACTTA 1849
617 hrValGlyProCysLeuIleAsnArgPheValAlaPheValArgLys 633
1850 CAGTTGGCGCTTGTAAATATAGTGTGTTGCCCTTTAGAGAAAGAGA 1899
634 ValSerAlaValGlnIleMetValLeuArgGlnGln 645
1900 GTGAGTGCAGTCCAGATCATGTACTTAGACAAACG 1935
seq_name: gb_v1:PEN288592
seq_documentation_block:
LOCUS PEN288592 2000 bp RNA VRL 29-AUG-2000
DEFINITION Porcine endogenous retrovirus PERV-B7 env gene for envelope
protein, genomic RNA.
ACCESSION AJ288592
VERSION AJ288592.1 GI:7414453
KEYWORDS env gene; envelope protein.
SOURCE porcine endogenous retrovirus.
ORGANISM porcine endogenous retrovirus

```


REFERENCE
 AUTHORS
 TITLE

Viruses; Retroid viruses; Retroviridae; Mammalian type C
 retroviruses: 1-Mammalian type C virus group.
 1 (bases 1 to 2000)
 Bosch, S., Arnold, C. and Jöstlin, A.
 Study of full-length porcine endogenous retrovirus genomes with

FEATURES	source
Location/Qualifiers	1. .2000
for name="Moraine endogenous retrovirus"	

gene
CDS

[illegible]

alignment_scores:	
Quality:	3329.00
Ratio:	5.243
Percent Similarity:	98.450
	Length: 645
	Gaps: 0
Percent Identity:	95.814

```
alignment_block:
```

US-09-171-553B-10 X PEN288592

Align seg 1/1 to: PEN288592 from: 1 to: 2000

1 MethHisProThrLeuSerTrpArgHisLeuProThrArgGlyGlyCysIupr 1
1 ATGCATCCACAGCTTAAGCTGGGGCCACCTTCGACGCTGGGGTGGAGAGCG 50
17 olysArgLeuArgIleIleProLeuSerPheAlaSerIleAlaTrpPheLeuTr 34
51 GAAGAAACAGGAAGAAATCCCTTAAGCTTGGCATCTTCCTCCGCCGCTTCTTA 100
34 hIeLeuThrIleThrProGlnAlaSerSerIysArgLeuIleAspSerSer 50
101 CTTCTAACAAATATTCGCCACGGCACTACTGTAAACGGCTATTATGACAGCTCG 150

51	AspProIisAngrProIeuSerProThrTrpPleuIleIleAsProkaph	67
151	AAACCCCATAGACCTTATCCCTTACCTGGCTGATATTAGACCCATATAC	200
67	rgIyValThrValAsnSerThrArgIyAlaIaProArgIyThrTrp	84
201	GGGIGTCACITGTAAATATAGCACTCGAGGTGTGCTCTAAGGACCTGT	250
84	rPrProGluIleuHisPheCysIleuArgIleuIleAsnProAlaValLysSer	100
251	GGCGTGAACGTCAATTCCTGGCTCCGATGTATTAACCCCGTGTAAAGC	300
101	ThrProAsnIleuValArgSerTyrGlyPheTyrCysCysPProGlyIynH	117
301	ACACCTCCCACTAGTCAGTCGTATGTATGGTCTATATGGTCCCAAGCAC	350
117	rgIuLysGluLysTyrCysGlyGlySerGlyGluSerPheCysArgArgT	134
351	AGAAAGAGAGAAATACGTGGGGTCTGGGGAATCCTTCTGTAGAGAT	400
134	rPserCysValThrSerAsnAspGlyAspPrpIysTrpProIleSerLeu	150
401	GGAGCTCCGTACCTCCACAGATGAGACGTGGAAATGGCGGATCTTCTC	450
151	GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrIy	167
451	CAGGACCGGGTAAATATCTCCTTTGTCAATTCGGCGCCGGGCAAGTACAG	500
167	sMeuLleLysIleuTyrIlyAspLysSerCysSerProSerAspLeuAspT	184
501	AGTATATAACTATATATAGATTAAGACTCTCTCCCACTACAGACTTACAT	550
184	YrIleuLysIleSerPheThrGluLysGlyLysGlnIuAsnIleGlnLys	200
551	ATCTAAAGATATAGTTTCACTGAAAGGAAACAGGAAATATTCAAAG	600
201	TrpIleAsnGlyMetSerTrpGlyIleValPheTyrIlySyrGlyGly	217
601	TGCTAAATGGTATGAGCTGGGGATGATTTTATTAATATGGCGGGGG	650
217	YAlGlySerThrIleuThrIleArgLeuAcqIleGlyThrGlyThrGluP	234
651	AGCAGGCTCCACTGATACCATTCGCCCTTAGATATGAGACCGGGGACAGAC	700
234	roProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProPro	250
701	CCCCGTGGCGAGTGGGAGCCCATTAAGTACGTGTGAACAGGGGCCCCCG	750
251	AlaIleuGluProProHisAsnIleuProValProGlnIleuThrSerLeuAr	267
751	GGCCTGGAGCCACCGCATTAACCTTGGCAGTGGCCCAATTAACTGGCGCG	800
267	qProAspIleThrGlnProProSerAsnSerThrThGlyLeuIleProT	284
801	GCCTGACATTAACACAGCGCGCTAGCAACGGTATCACCACTGATGTCTCTA	850
284	hrAsnThrProArgAsnSerProGlyValProValLysThrGlyGluArg	300
851	CCAAATATCTTAGAAGACCTCCAGGTGTCTCTTTAGACACGGAGAAAGGA	900
301	LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspPr	317
901	CTCTTCAGCTCATCTCAGGAGCTTCAAGCAATCAATCCACGAGACC	950
317	oAspAlaThrSerSerCysTrpPLeuCysLeuSerSerIlyProProIyTr	334
951	GGAGGCGCAATCTCTTGTGGCTTTGCTATCCCTCAAGGCGCTCCTATTT	1000
334	YrGlnGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArg	350
1001	ATGAGGGGATGGCTTAAGAGAGGAAATTCATGTGTACCAAGAGAGATGGA	1050
351	AsnGlnCysThrTrpGlySerArgAsnLysIleuThrIleuThrGluValSe	367

```

|||||
1051 AATCAATGTAACATGGGGGTCGCCGAATAAGCTTACCCCTCAGCAAGTTTC 1100
367 TGLYSGLYThrCysIleGlySAlaProProSerHisGlnHisLeu 384
|||||
1101 CGGGAAGGAACTGCATGGAAGAGCTCCCATCCCAACCAACCCCTT 1150
384 ystYSerThrValValTyrGluGlnAlaSerGlnAsnGlnTyrLeuVal 400
|||||
1151 GCATAGTACTGGTTTATGAAACAGCCCTCAGAAATACATGATTAGTA 1200
401 ProGlyTyrAsnArgTTrpTrpAlaCysAsnThrGlyLeuThrProCysVa 417
|||||
1201 CCTGGTTATACAGGTGGTGGCATGATCATCTGGTTAACCCCTGGT 1250
417 SerThrSerValPheAsnGlnSerLysAspLeuValMetValGlnI 434
|||||
1251 TTCACCTAGTCTTTAACCAATCCAAAGATTCTGTCTCATGGTCCAA 1300
434 LeuAlaProArgValTyrTyrHisProGluGluValValLeuAspGluTyr 450
|||||
1301 TCGTCCCGGAGTGTACTACCATCTCTGAGGAAGTGTCTGTATGATAT 1350
451 AspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAl 467
|||||
1351 GACTATCGGTAAACCGACCGAAAGAGAACCCGATCCCTTACCCTAGC 1400
467 aValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAla 484
|||||
1401 TGTATGCTCGGATTAGGAGCGCCCTTGGCGTAGGAACAGGACACTG 1450
484 LeuLeuIleThrGlyProGlnLeuGluLysGlyLeuGlyLeuHis 500
|||||
1451 CCTGATACAGACAGACAGCAGCTAGAGAAAGACTGGTGAAGCTATCAT 1500
501 AlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnLe 517
|||||
1501 GCGGCATGACAGAGATCTCCGAGCCTTAGAGAGCTGTAGCAACCT 1550
517 uGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 534
|||||
1551 AGAAGACTCCCTGACTCTTGTCTGAGAGTGGTTCTACCAACCGGAGG 1600
534 LysLeuAspLeuLeuPheLeuArgGluGlyLeuCysAlaAlaLeuLys 550
|||||
1601 GATTAGATCTGCTGTTCTAAGAGAGGTGGTATGTGTGACGCTTAAA 1650
551 GluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerNe 567
|||||
1651 GAAGAAATGTTGCTTCTATGTAGATCACTCAGAGCCATCAGACATCCAT 1700
567 tAsnLysLeuArgLysLysLeuGluArgArgArgArgLysGluAlaAla 584
|||||
1701 GACCAAGCTTAGAAGAGGTTAGAGAGCGCTCAAGAGGAAAGAGAGCTG 1750
584 spGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThrThr 600
|||||
1751 ACCAGGGGGGTTTGAAGGATGGTTCAACAGGCTCTGATGACACAC 1800
601 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuTh 617
|||||
1801 CTGCTTTCGCTCTGACGGGACCGGAGTAGTAGTCTGCTCTGTACTTAC 1850
617 rValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgy 634
|||||
1851 AGTTGGGCTTGCTTAATTAATAGGTTTGTCTTGTAGAGAAATGAG 1900
634 alSerAlaValGlnIleMetValLeuArgGln 645
|||||
1901 TAAGTGCAGTCCAGATCATGTACTTAGACAACGC 1935
seq_name: gb_v1: PEN288590

```

```

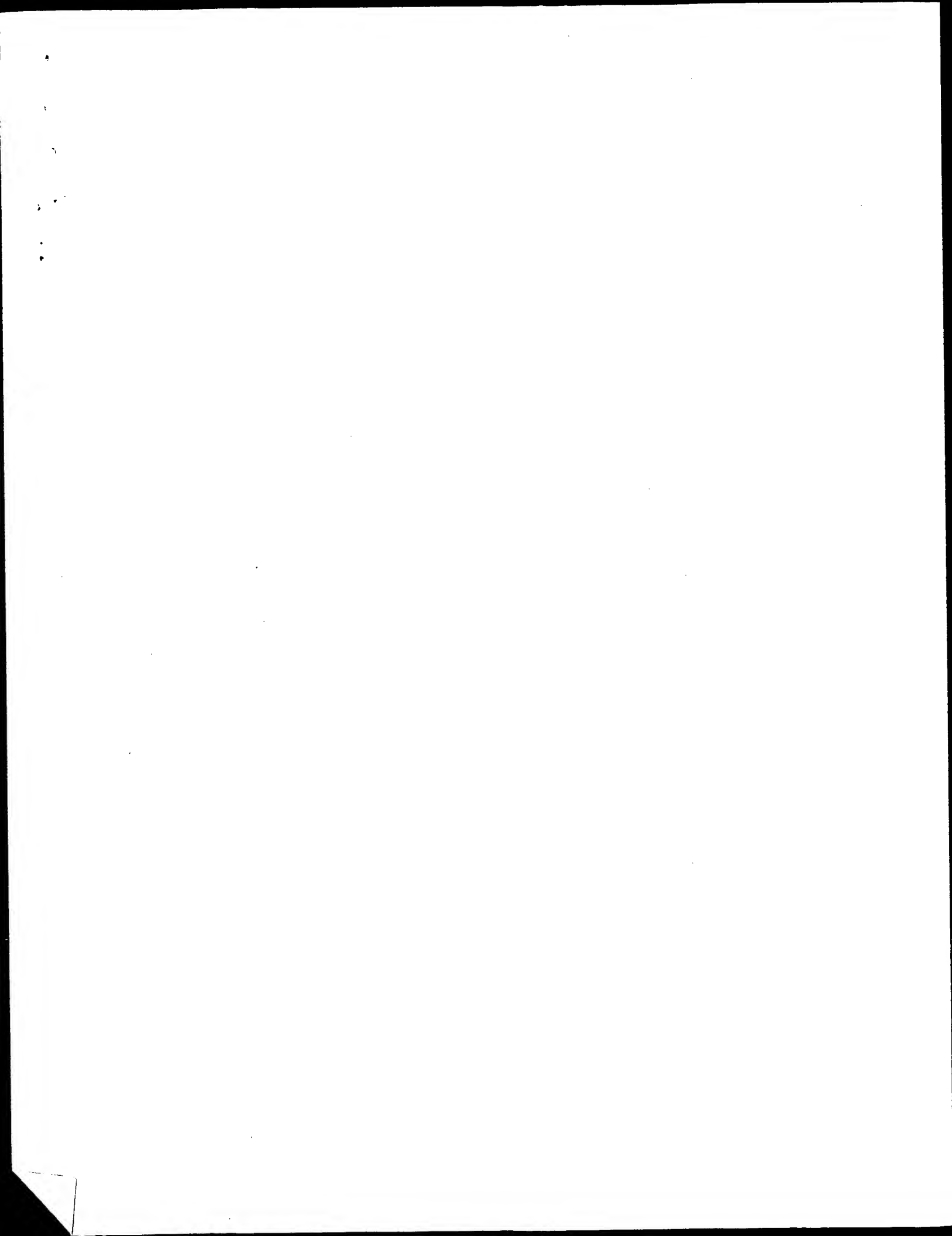
seq_documentation_block:
LOCUS      PEN288590      1996 bp      RNA      VRL      29-AUG-2000
DEFINITION Porcine endogenous retrovirus PERV-B4 env gene for envelope
ACCESSION  AJ288590
VERSION    AJ288590.1 GI:7414449
KEYWORDS   env gene; envelope protein.
SOURCE     porcine endogenous retrovirus.
ORGANISM   Porcine endogenous retrovirus
            Viruses; Retroviral viruses; Retroviridae; Mammalian type C
            retroviruses; 1-Mammalian type C virus group.
REFERENCE  1 (bases 1 to 1996)
AUTHORS   Bosch, S., Arnold, C. and Jestin, A.
TITLE     Study of full-length porcine endogenous retrovirus genomes with
            envelope gene polymorphism in a specific-pathogen-free large white
            swine herd
JOURNAL    J. Virol. 74 (18), 8575-8581 (2000)
MEDLINE    20411432
REFERENCE  2 (bases 1 to 1996)
AUTHORS   Bosch, S.
TITLE     Direct Submission
JOURNAL    Submitted (28-MAR-2000) Bosch S., Molecular Biology Department,
            AFSSA-Zoopole, Zoopole Les Croix B.P.53, Ploufragan 22440, FRANCE
FEATURES   Location/Qualifiers
            1..1996
             /organism="porcine endogenous retrovirus"
             /specific_host="Sus scrofa"
             /db_xref="taxon:51673"
             /clone="PERV-B4"
             /note="French Large White SPF-pig"
            1..1215
             /gene="env"
            1..1215
             /gene="env"
             /note="truncated, not functional"
             /codon_start=1
             /product="envelope protein"
             /protein_id="CAB6209.1"
             /db_xref="GI:7414450"
             /translation="MRPSSRRLLPTRGGEPRRLRIPASTAMFLTLYTPQASSK
            RLIDSSDHPRLSLITWLIIDPDGVYVNSRGAAPGTMPPELIFCLRLINPAVKSTP
            PNIVRSRPFCCCTGTEBERECGGRSFCGRMSCVTSNDGDMKPISTLQDRKFSVN
            SGREKRVMLYDKSCSPSDLDYKIFTEKROQNIQKWMGMSGIVFYKGGGA
            GSTLTTRILETHGTPVAVGPNIVLADEPAPLEPRLHPVQLTSLRPDITOPDN
            STTGLIPTNPRNSPGVPYTGRLFSLIQGAFOALNSTDPASSSCWCLSSGPPY
            EGMAREGKENVTRKHEHDOCTWCSRNKLTILEVSGKFTCIQKAPPSHQHLCNSTMYED
            ALENOYLVPYN"
BASE COUNT 521 a 499 c 503 g 472 t 1 others
ORIGIN
align_scores:
  Quality: 3130.50      Length: 646
  Ratio: 5.066          Gaps: 2
  Percent Similarity: 95.666      Percent Identity: 91.176
alignment_block:
US-09-171-553b-10 x PEN288590
Align seg 1/1 to: PEN288590 from: 1 to: 1996
1 MethisProthrLeuSerTTrpArgHisLeuProThrArgGlyGlyIupr 17
|||||
1 ATGGCTCCACAGTCACAGCCGCGCCACTCCGACTCGGGGTGAGAGCC 50
17 OlyArgLeuArgLLeProLeuSerPheAlaSerIleAlaTrpPheLeu 34
|||||
51 GAAGAAGCTAGGAATCCCTTAACTTCCCTCATCTGCTGGTTCCTTA 100
34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50
|||||
101 CTCTAACAGTGTACCCCGAGCGCAGTAGTAACGCTTATAGACAGCTCG 150

```


	seq_documentation_block:			VRL	29-AUG-2000		
LOCUS	PEN288591	2002 bp	RNA				
DEFINITION	Porcine endogenous retrovirus PERV-B6 env gene for envelope protein, genomic RNA.						
ACCESSION	AJ288591						
VERSION	AJ288591.1	GI:7414451					
KEYWORDS	env gene; envelope protein. porcine endogenous retrovirus. porcine endogenous retrovirus viruses; Retroid viruses; Retroviridae; Mammalian type C retroviruses; 1-Mammalian type C virus group.						
SOURCE	Bosch,S.; Arnault,C. and Jestin,A.						
ORGANISM	Study of full-length porcine endogenous retrovirus genomes with envelope gene polymorphism in a specific-pathogen-free large white swine herd						
REFERENCE	J. Virol.	74 (18),	8575-8581	(2000)			
AUTHORS	Medline	20411432					
TITLE	2 (bases 1 to 2002) Bosch,S.						
JOURNAL	Direct Submission						
MEDLINE	Submitted (28-Mar-2000) Bosch S., Molecular Biology Department,						
REFERENCES	AFSSA-Zoopole, Zoopole Les Croix B.P.53, Ploufragan 22440, FRANCE						
TITLE	Location/Qualifiers						
JOURNAL	1..2002						
MEDLINE	/organism="porcine endogenous retrovirus"						
AUTHORS	/specific_host="Sus scrofa"						
JOURNAL	/db_xref="taxon:61673"						
MEDLINE	/clone="PERV-B6"						
AUTHORS	/note="French Large White SPF-plg"						
JOURNAL	1..864						
MEDLINE	/gene="env"						
AUTHORS	1..864						
JOURNAL	/gene="env"						
MEDLINE	/note="truncated, not functional"						
AUTHORS	/codon_start=1						
JOURNAL	/product="envelope protein"						
MEDLINE	/protein_id="CAB86210.1"						
AUTHORS	/db_xref="GI:7414452"						
JOURNAL	/translation="MGPRRSSRRLPNMGGERKLRPLSPASIAMFLTFNSDPOASSKR						
MEDLINE	RLDISDPHRPLSLTLWLIIIDPDITGVYNSRVGVPAGTWMPDLHFCRLINPAVKSTPP						
AUTHORS	PNIIVRSYGFCPCETKEEYGGSGRESFGMSVCTNSDGPMKWPIILODRKRFNYPNV						
JOURNAL	SGLWIKGYFMELTKDKSCSPSDIOLTKISPLEKGQEMTIOWKNIMSWGITYFYKKGGGA						
MEDLINE	GSTLTIRLRETGTGEPRPVAVGNPVIILAEGCPALPEPHNLVPOLITSILRPDIIPPDN						
AUTHORS	STTGLISTDYHA"						
BASE COUNT	531 a	495 c	503 g	472 t	1 others		
ORIGIN							
<hr/>							
alignment_scores:							
Quality:	3056.50	Length:	646				
Ratio:	4.978	Gaps:	2				
Percent Similarity:	95.046	Percent Identity:	90.402				
<hr/>							
Alignment block: US-09-171-553B-10 x PEN288591 ..							
Align seg 1/1 to: PEN288591 from: 1 to: 2002							
1 MethisProThrLeuSerTrpArgHisLeuProthrArgGlyGluIupr 17							
1 ATGGGATCCCAAGTCACAAGCGCGGCCAACCTCCGAAATTGGGGTGAGAACC 50							
17 OlysArGlueAtArgLieProleuSerSheAlaSerlieAlatrpheLeut 34							
51 GAAAGATTGAGAACCCCCCTTAAGCTTGCGCTTCGCATCGGCTGTTCCTTA 100							
34 htleuthrliehprrcginAlaSerSerLyArgluelileaspserser 50							
101 CTTTAAAGTAGATCCCAAGGCCAGTAGAAACGGCTTAATAACAGACTTG 150							
51 AsnPrOHsarGrProleuSerProthrTrpHeulleAspprotasph 67							

151 GACCCCATGAGACCCCTTATCCCTACGTGGCTGATTATGACCCGTATAC 200
67 rglYvalThvValasnSerThrArgLysAlaIaProArgLysThrTrp 84
201 GGGTGTACACTGTTAAATGAGCTAGAGGTGTGGCTCTAGAGCACTGGT 250
84 rPrpGluuLeuHISpHeCyLeuArgLeuIleIasnProAlaValLysSer 100
251 GGCTGTGACGCACTTCTGGCTCCGATTGATTAAACCCCTCTGTTAAAC 300
101 ThrProAsnLeuValAcSerTyrGlyPheTyrCySAspProGlyIyn 117
301 ACACCTCCCACTAGTCCTCACTATTATGGTTCTATTGCTGCCAGGCAC 350
117 rGluLysGluLysTyrCySgLYserGlyLysSerPheCyArgArgT 134
351 AGAGAAAGAGAGATCTAGTGGGGGTCCAGGAAATCCTCTGAGGAGAT 400
134 rPserCySValIlnrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150
401 GGAGCTGGCGCACCTCCAAAGATGAGACTGTGAAAGGCCGATCTCTTC 450
151 GlAspArgValLysPheSerPheValasnSerGlyProGlyLysTyrIy 167
451 CAGAGCCGGTAAATTCCTCTTTGCAATTCGGCCCGGCGCAAGTACAA 500
167 smetMetLysLeuTyrLysAspLysSerCySserProSerAspLeuAspT 184
501 AGTGAATGAGCACTATTAAAGATTAAGAGCTGTGCCCATCAGACTTAA 550
184 yrlLeuLysIleSerPheThrGluLysGlyLysGlnIuasnIleGlnLys 200
551 ATCTAAAGATTAACCTTCACTGAAAAAGAAACAGAGAAATATTCAAAAG 600
201 TrpIleAsnGlyMetSerTrpGlyIleValPheTyrLysTyrGlyGly 217
601 TGGATTAATATGTAATGAGCTGGGGAATAGTTTATTAAATATGAGAGGG 650
217 yAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyTrpGlu 234
651 AGCAGGGTCCACTTAAACCATTCGGCTTAGATAGAGAGGGAGCAAC 700
234 rOpProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyPro 250
701 CCCCTGTGGCAGAGGGACCAACATAGTACTGGCGAAGAGGGGCCCG 750
251 AlaLeuGluProProIleAsnLeuProValProGlnLeuThrSerLeu 267
751 GCCCTGAGCACACACACACTTGGCGGTGCCCGCAAGTTAACTCGCTGG 800
267 gProAspIleThrGlnProProSerAsnSerThrThrGlyLeuIlePro 283
801 GCCTGACATTAACACAGCCCGCTGACACAGTACCACTGGATTGATTC 850
284 ThrAsnThrProArgAsnSerProGlyValProValLysThrGlyGln 300
851 ACCAAACACCCAGAAACCCCGGTGTCTCTGTAAACAGGACAAG 900
300 GluPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerTrpAsp 317
901 ACCTCTACAGCTCATTCACAGGAGCTTCCAAAGCCATCACTCCAGCAG 950
317 rAspAlaThrSerSerCysTrpLeuCySLeuSerSerGlyProProTrp 333
951 CGATGCCACTCTCTTATTGGCTTGTGATCTCTCAAGGGCTCTTAT 1000
334 TyrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHis 350
1001 TATGAGGGAATGGCTAGAGAAAGGAATTCATCATGTGACCAAGACAT 1050
350 gAsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluVal 367
1051 AAATCAATGATCAATGGGGGTCCCGAAATAGCTCAACCTCAACCGAAAT 1100

```
367  erGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeu 383
|||||
1101  CCGGGAAGGGGACATGCATAGAAAGGTCCTCCCATCCACCAACACCTT 1150
|||||
384  CysTyrSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeuVa 400
|||||
1151  TGCATAGTACTAGGTTTAGACAGCCCTTAGAAATCAGTATTAGT 1200
|||||
400  LProGlyTyrAsnArgTyrPheAlaCysAsnThrGlyLeuThrProCysV 417
|||||
1201  ACCGTGTTATACGATGATGGTGGCATCATCTGGGTAACTCCCTGTG 1250
|||||
417  aSerThrSerValPheAsnGlnSerLysAspLeuCysValMetValGln 433
|||||
1251  TTTCTACACAGAGCTCTTAACCAATCCAAAGATTCTG .GTCATGCTCCAG 1299
|||||
434  IleValProArgValTyrTyrHisProGluGluValValLeuAspGlnTyr 450
|||||
1300  CTGCTCCCTGGGTACACTACCTCCAGAGAAANTCATCGATGAGTA 1349
|||||
450  rAspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuA 467
|||||
1350  GGACTATCGGCTAACCCGCTCAAAAAGAGACCCGTGACCTTAACCTAG 1399
|||||
467  IValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAla 483
|||||
1400  CCGTTATACTAGGATTAGGATGGCT...GGTGTGGGACAGGACAGCT 1446
|||||
484  AlaLeuIleThrGlyProGlnGlnLeuGlyLysGlyLeuGlyLeuH 500
|||||
1447  GCCCTGATCACAGACACAGCAGATTAGAAAAAGGACTGGTGAGCTACA 1496
|||||
500  sAlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsn 517
|||||
1497  TCGCGCATTAACGAAGATCTCCGAGCCTTAGAGAAATCTGTTAACAA 1546
|||||
517  euGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 533
|||||
1547  TAGAAGAAATCCCGACTCT.TTGTCTGAAGTGGTCTACAGAACTGAGG 1595
|||||
534  GlyLeuAspLeuLeuPheLeuArgGluGlyLeuCysAlaAlaLeuLys 550
|||||
1596  GGATTAGATCTGCTGTTCTAAAAGAGGTGGTTATGTGACGCCCTTAG 1645
|||||
550  sGluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSer 567
|||||
1646  AGAAGAATGTGCTTATGTGATCACTCAGAGAGCATCAGAGACTCCA 1695
|||||
567. eTAsnLysLeuArgLysLysLeuGluArgArgArgGluArgGluAla 583
|||||
1696  TCGAGCAAGCTTAGAGAAAGTTAGAGACGTCGAAGGGAAGAGAGGCT 1745
|||||
584  AspGlnGlyTyrPheGluGlyTyrPheAsnArgSerProTyrMetThrTh 600
|||||
1746  GACCAAGGGGTGTTGAGAGATGTTCAACAGGCTCTCTTGATGATCAC 1795
|||||
600  rLeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuT 617
|||||
1796  CCTGCTCTCTGCTGATGGACCCCTAGTGTCTGCTCTGCTGCTGTTA 1845
|||||
617  hrValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArg 633
|||||
1846  CAGTTGGGCTTGCCTTAATTAATAGTTATGCTTGTAGAGAAATGA 1895
|||||
634  ValSerAlaValAlaGlnIleMetValLeuArgGln 645
|||||
1896  GTGAGTGCAGTCCAGATCATGTACTTAGACACAGC 1931
|||||
```



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2002, 02:56:21 ; Search time 10411.9 Seconds
(without alignments)
28.520 Million cell updates/sec

Title: US-09-171-553B-14

Perfect score: 18

Sequence: 1 ccacagtcgtacaccacy 18

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hgt.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_om.*

20: em_or.*

21: em_ov.*

22: em_pat.*

23: em_ph.*

24: em_pl.*

25: em_ro.*

26: em_sts.*

27: em_sy.*

28: em_un.*

29: em_vi.*

30: em_htgo_hum.*

31: em_htgo_inv.*

32: em_htgo_rod.*

33: em_htg_hum.*

34: em_htg_inv.*

35: em_htg_rod.*

36: em_htg_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

8

Result No.	Score	Query Match	Length	DB ID	Description
C 1	18	100.0	590	14	PEN298074
C 2	18	100.0	633	6	AX052622
C 3	18	100.0	704	6	AX052621
C 4	18	100.0	746	14	PEN298073
C 5	18	100.0	1042	14	PERV17012
C 6	18	100.0	3025	4	AF147808
C 7	18	100.0	3482	6	AX002804
C 8	18	100.0	3482	14	PERENV2
C 9	18	100.0	4918	6	AX052637
C 10	18	100.0	6076	6	AX052636
C 11	18	100.0	6076	6	AX052647
C 12	18	100.0	7333	4	AF038601
C 13	18	100.0	7333	6	AR130474
C 14	18	100.0	7362	6	AX052634
C 15	18	100.0	7873	6	AX052638
C 16	18	100.0	8060	6	AR130473
C 17	18	100.0	8132	4	AF038599
C 18	18	100.0	8132	4	AF038600
C 19	18	100.0	8132	6	AR130475
C 20	18	100.0	8196	6	A66552
C 21	18	100.0	8209	6	A66553
C 22	18	100.0	8750	14	PEN133818
C 23	18	100.0	8763	14	PEN293657
C 24	18	100.0	8849	14	PEN133817
C 25	18	100.0	8918	14	PEN133816
C 26	17	94.4	18	6	AX118774
C 27	17	94.4	20	6	AR130533
C 28	17	94.4	28	6	AX002816
C 29	16.4	91.1	595	14	PEN298075
C 30	16.4	91.1	610	4	AF286463
C 31	16.4	91.1	1387	10	AF161301
C 32	16.4	91.1	1491	10	AF106919
C 33	16.4	91.1	1878	3	POV278955
C 34	16.4	91.1	2945	3	AF177661
C 35	16.4	91.1	8918	14	PEN293656
C 36	16.4	91.1	90251	2	AC074230
C 37	16.4	91.1	110000	2	AC073804
C 38	16.4	91.1	110000	2	AC087331
C 39	16.4	91.1	110000	2	AC087782
C 40	16.4	91.1	110000	2	AC090843
C 41	16.4	91.1	123835	10	MM303011
C 42	16.4	91.1	148588	2	AC073827
C 43	16.4	91.1	165563	2	AC083752
C 44	16.4	91.1	170906	2	AC084314
C 45	16.4	91.1	174533	2	AC073718

ALIGNMENTS

RESULT	1	Pen298074/c
LOCUS	PEN298074	Porcine endogenous retrovirus type C proviral LTR sequence (strain PERV-B(43)), 590 bp.
DEFINITION	PEN298074	590 bp DNA
ACCESSION	AJ298074	VRL
VERSION	AJ298074.1	17-JUL-2001
KEYWORDS	long terminal repeat; LTR.	
SOURCE	porcine endogenous retrovirus.	
ORGANISM	porcine endogenous retrovirus	
REFERENCE	Viruses; Retroviral viruses; Retroviridae; Mammalian type C retroviruses; 1-Mammalian type C virus group.	
AUTHORS	1 (bases 1 to 590)	
TITLE	Czauderna,F., Fischer,N., Boller,K., Kurth,R. and Tonjes,R.R.	
JOURNAL	Establishment and characterization of molecular clones of porcine endogenous retroviruses replicating on human cells	
MEDLINE	J. Virol. 74 (9), 4028-4038 (2000)	
REFERENCE	20219394	
AUTHORS	2 (bases 1 to 590)	
	Scheef,G., Fischer,N., Krach,U. and Tonjes,R.R.	


```

/db_xref="taxon:61673"
LTR 1..746
misc_feature 1..584
TATA_signal /note="U3 region"
misc_feature 556..561
misc_feature 585
misc_feature /note="Cap site"
misc_feature 585..663
misc_feature /note="R region"
misc_feature 664..746
misc_feature /note="U5 region"
BASE COUNT 182 a 186 c 184 g 194 t
ORIGIN
Query Match 100.0%; Score 18; DB 14; Length 746;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ccacagtcgtacaccacg 18
|||||
Db 622 CCACAGTCGTACACCACG 605

RESULT 5
PERY17012/c
LOCUS PERY17012 1042 bp DNA VRL 14-APR-2000
DEFINITION Porcine endogenous retrovirus DNA for long terminal repeat.
ACCESSION Y17012
VERSION Y17012.1 GI:3005984
KEYWORDS long terminal repeat; LTR.
SOURCE porcine endogenous retrovirus.
ORGANISM Porcine endogenous retrovirus.
REFERENCE 1 (bases 1 to 1042)
AUTHORS Czauderna,F.; Fischer,N., Boller,K., Kurth,R. and Tonjes,R.R.
TITLE Establishment and characterization of molecular clones of porcine
endogenous retroviruses replicating on human cells
J. Virol. 74 (9), 4028-4038 (2000)
20219394
REFERENCE 2 (bases 1 to 1042)
AUTHORS Tonjes,R.R.
TITLE Direct Submission
JOURNAL Submitted (30-MAR-1998) R.R. Tonjes, Paul-Ehrlich Institut,
Paul-Ehrlich Str. 51-59, P.O.Box 1740, D-63207 Langen, FRG
LOCATION/Qualifiers
FEATURES
source
1..1042
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
/tissue_type="kidney"
/cell_line="PK15"
1..1042
LTR 214 a 270 c 289 g 269 t
BASE COUNT
ORIGIN
Query Match 100.0%; Score 18; DB 14; Length 1042;
Best Local Similarity 100.0%; Pred. No. 9.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ccacagtcgtacaccacg 18
|||||
Db 503 CCACAGTCGTACACCACG 486

RESULT 6
AF147808/c
LOCUS AF147808 3025 bp DNA MAM 27-APR-2000
DEFINITION Sus scrofa domestica porcine endogenous retrovirus type C gag-pol
precursor, gene, partial cds.
ACCESSION AF147808
VERSION AF147808.1 GI:7650139

/db_xref="taxon:61673"
LTR 1..746
misc_feature 1..584
TATA_signal /note="U3 region"
misc_feature 556..561
misc_feature 585
misc_feature /note="Cap site"
misc_feature 585..663
misc_feature /note="R region"
misc_feature 664..746
misc_feature /note="U5 region"
BASE COUNT 182 a 186 c 184 g 194 t
ORIGIN
Query Match 100.0%; Score 18; DB 4; Length 3025;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ccacagtcgtacaccacg 18
|||||
Db 461 CCACAGTCGTACACCACG 444

RESULT 7
AX002804/c
LOCUS AX002804 3482 bp DNA PAT 21-AUG-2000
DEFINITION Sequence 3 from Patent WO9853104.
ACCESSION AX002804
VERSION AX002804.1 GI:9885132,
KEYWORDS porcine endogenous retrovirus.

```

```

KEYWORDS domestic pig.
SOURCE Sus scrofa domestica
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 3025)
AUTHORS Blusch,J.H., Seelmeir,S. and von der Helm,K.
TITLE Evidence for extreme sequence homogeneity of the PERV A/B 5'
LTR/leader/gag/prot region by predicted targeted amplification from
a Du x BL pig
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3025)
AUTHORS Blusch,J.H.
TITLE Direct Submission
JOURNAL Submitted (03-MAY-1999) Virology, Max-von-Pettenkofer-Institute,
Pettenkofer-Str. 9a, Munich 80336, Germany
FEATURES
source
1..3025
/organism="Sus scrofa domestica"
/strain="breed Duroc x German Landbreed, Du x DL3791"
/sub_species="domestica"
/db_xref="taxon:9825"
1..3025
repeat_region
1..3025
/note="PERV A/B; porcine endogenous retrovirus type C"
/rpt_family="perv"
/rpt_type="dispersed"
587..605
/note="primer binding site for tRNA-Gly"
1032..>3025
/note="gag protein produced by proteolytic cleavage of the
mature protease out of the precursor"
/codon_start=1
/transl_except="(pos:2604..2606,aa:Gln)
/product="gag-pol precursor"
/protein_id="AAF65926.1"
/db_xref="GI:7650140"
/translation="MGQTVTTPLSLTLDHWTEVRSRAHNSLVQVKKGPWOTFCASEWP
TFDVGWPSQGTNSEIILAVKALIFOTGSPSHDQEPYILTWDLAEDPPVWKPLN
KPRKPGRIILALGEKNKSAEKVSPRIYPLEETWPEQVPVPPYPAGGAVRG
PSAPGAPVVEGPAAGTRSRRCATPERTDEIILPTVGPMPGQQLQLOIWPFS
ADLYNWKTHPPFSDQRLTGLVLSHQTDDCQQLQTLFTTEERILLEA
RKNVPGADGRTQLQNEIDMGFPLTRPDYNTAEGRESKIYRQALVAGLRGARRP
TKLAKRYVMQGNPPSVFLERLMEAFRRFPDPTSEAKASVALAFIGOSALDIR
KKLQRLGLEQAEALRDVREAKVYVRETEKEQKEREEREERDRQENLT
KILAAVVEGKSSRERERDFKIRSGPROSGNLGNTPLDKOCAYCKEKGHWARNCPK
KGNKGKVLAEEDKQGRGSDPLPEPRVTLKVEGQPVFVLDVGAHSVLPLQPKG
LKEKSWMGATGQRYPMTRTRTDVLDGVRVTHSLVIPCVPVLLGLDLTKWGAQ
ISFEQGRPEVSVNNKPTITVLTQLDDEYRLYSPQVKPDQD"
1032..2603
mat_peptide /product="gag protein"
2604..>3025
mat_peptide /product="aspartic protease"
BASE COUNT 743 a 749 c 849 g 684 t
ORIGIN
Query Match 100.0%; Score 18; DB 4; Length 3025;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ccacagtcgtacaccacg 18
|||||
Db 461 CCACAGTCGTACACCACG 444

RESULT 7
AX002804/c
LOCUS AX002804 3482 bp DNA PAT 21-AUG-2000
DEFINITION Sequence 3 from Patent WO9853104.
ACCESSION AX002804
VERSION AX002804.1 GI:9885132,
KEYWORDS porcine endogenous retrovirus.

```

ORGANISM porcine endogenous retrovirus
viruses; Retroviridae; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.

REFERENCE 1 (bases 1 to 3482)
AUTHORS Stoye,J.P. and Weiss,R.A.
TITLE Detection of retroviral subtypes based upon envelope specific sequences

JOURNAL Patent: WO 9853104-A 3 26-NOV-1998;
MEDICAL RES COUNCIL (GB); STOYE JONATHAN PAUL (GB)

FEATURES
source
1..3482
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"

BASE COUNT 927 a 854 c 867 g 834 t
ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 3482;
Best Local Similarity 100.0%; Pred. No. 8.4;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ccacagtcgtacaccag 18
|||||

Db 3439 CCACAGTCGTACACCAG 3422

RESULT 8
PERENV2/c
LOCUS PERENV2 3482 bp RNA VRL 14-OCT-1997
DEFINITION Porcine endogenous retrovirus env gene, 3482 bp.
ACCESSION Y12239
VERSION Y12239.1 GI:2576324
KEYWORDS env gene.
SOURCE porcine endogenous retrovirus.
ORGANISM porcine endogenous retrovirus
Viruses; Retroviridae; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.

REFERENCE 1 (bases 1 to 3482)
AUTHORS Le Tissier,P., Stoye,J.P., Takeuchi,Y., Patience,C. and Weiss,R.A.
TITLE Two sets of human-tropic pig retrovirus
JOURNAL Nature 389 (6652), 681-682 (1997)
MEDLINE 97478526
REFERENCE 2 (bases 1 to 3482)
AUTHORS Stoye,J.P.
TITLE Direct Submission
JOURNAL Submitted (02-APR-1997) J.P. Stoye, National Institute for Medical
Research, Virology, The Ridgeway Mill Hill, London, NW7 1AA, UK
REMARK revised by author 16-SEP-97

FEATURES
source
1..3482
/organism="porcine endogenous retrovirus"
/specific_host="Sus scrofa"
/db_xref="taxon:61673"
/note="PERV-B"
911..2884
/gene="env"
911..2884
/gene="env"
/codon_start=1
/protein_id="CAA72928.1"
/db_xref="GI:2576325"
/db_xref="SPRMBL:041173"
/translation="MHPTLSWRHLPTRGGEKPKRLRPLSFASIAFWLTITTPQASSK
RLDSSNPHRPLSLTLLIDPDGTGVNSTRGVAPRGTWMPHLFCLRLINPAVKSTP
PNLVRSGYFCPCTEKEKYGCGSGESFCRRSCVTSDGDMKWPISLODRVKSFTV
SGPKYKMKLYKDKSCSPDLDYKLSITEKGKQENIQKINGMSGIVFYKGGGA
GSULTRLRETGETPVPVAGPKVLAEGPPALPEPHNLVPQLSLRPDIQPPSN
GTGLIPTPRNSPGVPVKTQRLFLIQGAQAINSTDPPATSSCWLCLSGGPPY
AGMAKEGFNVTREHNRQCTWGSNKLTLEVSQKTCIGKAPPSQHILCYSTVVEQ
ASENQYLPGYNRWACNTGLTPTCVSTVFNSQKDFCVMWIVPVYVHPEAVTVEY
DYRNRKRPVSLTLAVMLGLQTAVGVTGTAALITGPQLEKGLHGAHGAEDLR
ALRESVNLSESLTSLSEVLQNRGLDLLFLREGGLCAALKKECCFYVDHSGAIFDS
MSKLRERLRRRRERREADQWFEQFNRSFPMWMTLLSALTGPLVLLILLITVGPCLIN

BASE COUNT 927 a 854 c 867 g 834 t
ORIGIN

Query Match 100.0%; Score 18; DB 14; Length 3482;
Best Local Similarity 100.0%; Pred. No. 8.4;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ccacagtcgtacaccag 18
|||||

Db 3439 CCACAGTCGTACACCAG 3422

RESULT 9
AX052637/c
LOCUS AX052637 4918 bp DNA PAT 12-JAN-2001
DEFINITION Sequence 22 from Patent WO0071726.
ACCESSION AX052637
VERSION AX052637.1 GI:12226827
KEYWORDS
SOURCE porcine endogenous retrovirus.
ORGANISM porcine endogenous retrovirus
Viruses; Retroviridae; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.

REFERENCE 1 (bases 1 to 4918)
AUTHORS Federspiel,M.J.
TITLE Methods to inhibit infectious agent transmission during
xenotransplantation
JOURNAL Patent: WO 0071726-A 22 30-NOV-2000;
Mayo Medical Ventures (US)
FEATURES
source
1..4918
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"

BASE COUNT 1382 a 1173 c 1183 g 1180 t
ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 4918;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ccacagtcgtacaccag 18
|||||

Db 4781 CCACAGTCGTACACCAG 4764

RESULT 10
AX052636/c
LOCUS AX052636 6076 bp DNA PAT 12-JAN-2001
DEFINITION Sequence 21 from Patent WO0071726.
ACCESSION AX052636
VERSION AX052636.1 GI:12226826
KEYWORDS
SOURCE porcine endogenous retrovirus.
ORGANISM porcine endogenous retrovirus
Viruses; Retroviridae; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.

REFERENCE 1 (bases 1 to 6076)
AUTHORS Federspiel,M.J.
TITLE Methods to inhibit infectious agent transmission during
xenotransplantation
JOURNAL Patent: WO 0071726-A 21 30-NOV-2000;
Mayo Medical Ventures (US)
FEATURES
source
1..6076
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"

BASE COUNT 1613 a 1512 c 1609 g 1342 t
ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 6076;
 Best Local Similarity 100.0%; Pred. No. 7.9;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ccacagtcgtacaccag 18
 |||||
 Db 617 CCACAGTCGTACACCAG 600

RESULT 11
 AX052647/c
 LOCUS AX052647 6076 bp DNA PAT 12-JAN-2001
 DEFINITION Sequence 32 from Patent WO0071726.
 ACCESSION AX052647
 VERSION AX052647.1 GI:12226837
 KEYWORDS
 SOURCE porcine endogenous retrovirus.
 ORGANISM porcine endogenous retrovirus

REFERENCE 1 (bases 1 to 6076)
 AUTHORS Federspiel, M.J.
 TITLE Methods to inhibit infectious agent transmission during
 xenotransplantation
 JOURNAL Patent: WO 0071726-A 32 30-NOV-2000;
 Mayo Medical Ventures (US)

FEATURES
 source
 1. .6076
 /organism="porcine endogenous retrovirus"
 /db_xref="taxon:61673"
 BASE COUNT 1619 a 1505 c 1598 g 1354 t
 ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 6076;
 Best Local Similarity 100.0%; Pred. No. 7.9;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ccacagtcgtacaccag 18
 |||||
 Db 617 CCACAGTCGTACACCAG 600

RESULT 12
 AF038601/c
 LOCUS AF038601 7333 bp mRNA MAM 21-MAY-1998
 DEFINITION Sus scrofa porcine endogenous retrovirus ERV-PK15 mRNA, complete
 sequence.

ACCESSION AF038601
 VERSION AF038601.1 GI:3133304
 KEYWORDS
 SOURCE pig.
 ORGANISM Sus scrofa

REFERENCE 1 (bases 1 to 7333)
 AUTHORS Akiyoshi, D.E., Denaro, M., Zhu, H., Greenstein, J.L., Banerjee, P. and
 Fishman, J.A.
 TITLE Identification of a full-length cDNA for an endogenous retrovirus
 of miniature swine

J. Virol. 72 (5), 4503-4507 (1998)
 MEDLINE 98216827
 REFERENCE 2 (bases 1 to 7333)
 AUTHORS Fishman, J.A.
 TITLE Direct Submission

JOURNAL Submitted (16-DEC-1997) Infectious Disease Unit, Massachusetts
 General Hospital, 55 Fruit Street, Boston, MA 02114, USA
 REFERENCE 3 (bases 1 to 7333)
 AUTHORS Fishman, J.A.
 TITLE Direct Submission

JOURNAL Submitted (15-MAY-1998) Infectious Disease Unit, Massachusetts
 General Hospital, 55 Fruit Street, Boston, MA 02114, USA
 REMARK Sequence update by submitter

COMMENT On May 16, 1998 this sequence version replaced gi:3116445.
 FEATURES
 source

1. .7333

/organism="Sus scrofa"
 /strain="miniature swine"
 /db_xref="taxon:9823"
 /cell_line="PK-15"

/note="porcine endogenous retrovirus ERV-PK15; infective
 virus but represents a naturally occurring transcript"

1. .578

598. .2172

/codon_start=1

/product="gag protein"

/protein_id="AAC16766.1"

/db_xref="GI:3116446"

/translation="MGQTVTTPLSLTDHWTVEVRAHNSLVQVKGQWTFCASEWP
 TFDVGPSEGTNSEIILAVKAIIFQTGPGSHPDQEPYILTWDLAEDPPWPKPWLN
 KPRKGPRIALGKKNHSAEKVPSRIYPIEIEPTWPEQVPVPPYPAQCAVGR
 PSAPGAPVVECPAGATSRGATPERTDEIALLPLRTYGPMPGGQLQLOYWPFSS
 ADLYNWKTHHPFSDPDRGLTGLVESLMFSHQPTWDDCCQLQTLFTTEERILLLEA
 KKNVPGADGRPTQONEIDMGFPPLTPGWDYNTAEGRESLKIYQALVAGLGRARRP
 TNLAKREVMPGNEPVSFLELMEAFREFPTPDPTSEAKASVALAFIQSALDIR
 KKLQLEGLQAEILDLVREAEKVYIRRETEKEKEKEKEKEKEKEKEKEKEKILIT
 KILAAVVEGKSSRRERDRFKIRSGPROSGNLGNRTPLDKDQCAYCKEKGHWARNCKP
 KGNKGPVLALEEDKD"

2320. .3552

/note="truncated protein"

/codon_start=1

/product="pol protein"

/protein_id="AAC16767.1"

/db_xref="GI:3133305"

/translation="MGATGQRQYPTTTRTVDLGVGRVTHSVFLVIPCVPPLIGRDL
 TKGAQISFEGREPEVNNPITVLTLDDEYRLSPQVKPDQDQISWLEQFPQAM
 AETAGMLAKQVPPQVQLKASATPSVQVPLSREAREGHWPHVRLIQGLILPVQ
 SPWNTPLPVKPGTNDYRPVQDLREVNKRQVDIHTVPNPVNLNLSALPERNWTVL
 DLKDAFFCLRLHPTSQLFTPEWRDPTGRTGQLTWTLPQGFKNSTPTIDEALHRL
 ANRIHQPVQVTLLOYVDDLLAGATKODCLGKTALLLESLDLGYRASAKKQICRE
 VTYLGYSLRGQRWLTEARKTVVQIIPAPTAKOVREFLGTAGFCRLWIFGFATLAAP
 LYPLTKKGCGCPQQGK"

4803. .6725

/note="env protein; N terminus deleted"

/codon_start=1

/pseudo

/base_count

BASE COUNT 1984 a 1788 c 1885 g 1676 t

ORIGIN

Query Match 100.0%; Score 18; DB 4; Length 7333;
 Best Local Similarity 100.0%; Pred. No. 7.8;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ccacagtcgtacaccag 18
 |||||
 Db 7279 CCACAGTCGTACACCAG 7262

RESULT 13

ARI30474/c

LOCUS ARI30474 7333 bp DNA PAT 16-MAY-2001

DEFINITION Sequence 2 from patent US 6190861.

ACCESSION ARI30474

VERSION ARI30474.1 GI:14118799

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 7333)

AUTHORS Fishman, J.A.

TITLE Molecular sequences of swine retroviruses method of using

JOURNAL Patent: US 6190861-A 2 20-FEB-2001;

FEATURES Location/Qualifiers

1. .7333

source

BASE COUNT 1984 a 1788 c 1885 g 1676 t
ORIGIN /organism="unknown"

Query Match 100.0%; Score 18; DB 6; Length 7333;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ccacagtcgtacaccag 18
|||||
Db 7279 CCACAGTCGTACACCAG 7262

RESULT 14
AX052634/c 7362 bp DNA PAT 12-JAN-2001
LOCUS
DEFINITION Sequence 19 from Patent WO0071726.
ACCESSION AX052634
VERSION AX052634.1 GI:12226824

KEYWORDS porcine endogenous retrovirus.
SOURCE porcine endogenous retrovirus
ORGANISM Viruses; Retrovirus; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
1 (bases 1 to 7362)
AUTHORS Federspiel, M.J.
TITLE Methods to inhibit infectious agent transmission during
xenotransplantation
JOURNAL Patent: WO 0071726-A 19 30-NOV-2000;
Mayo Medical Ventures (US)

FEATURES
source 1. 7362
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"

BASE COUNT 1997 a 1821 c 1881 g 1663 t
ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 7362;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ccacagtcgtacaccag 18
|||||
Db 682 CCACAGTCGTACACCAG 665

RESULT 15
AX052638/c 7873 bp DNA PAT 12-JAN-2001
LOCUS
DEFINITION Sequence 23 from Patent WO0071726.
ACCESSION AX052638
VERSION AX052638.1 GI:12226828

KEYWORDS porcine endogenous retrovirus.
SOURCE porcine endogenous retrovirus
ORGANISM Viruses; Retrovirus; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
1 (bases 1 to 7873)
AUTHORS Federspiel, M.J.
TITLE Methods to inhibit infectious agent transmission during
xenotransplantation
JOURNAL Patent: WO 0071726-A 23 30-NOV-2000;
Mayo Medical Ventures (US)

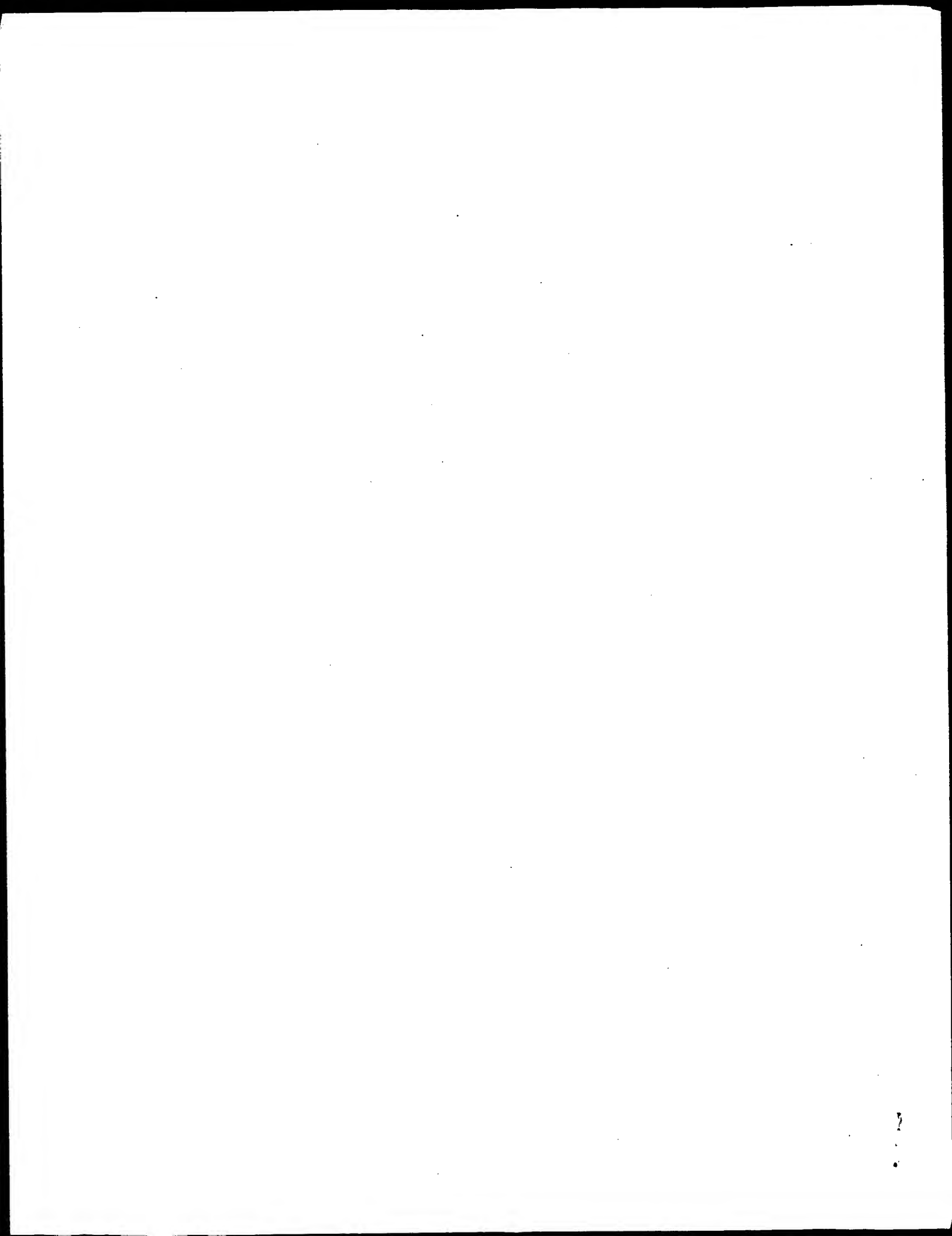
FEATURES
source 1. 7873
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"

BASE COUNT 2200 a 1914 c 1961 g 1798 t
ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 7873;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ccacagtcgtacaccag 18
|||||
Db 7708 CCACAGTCGTACACCAG 7691

Search completed: February 24, 2002, 02:56:23
Job time: 21294 sec



○

A66551	3320 bp	DNA	PAT	29-MAR-1999
LOCUS	Sequence 1 from Patent WO9740167.			
DEFINITION	A66551			
ACCESSION	A66551.1			
VERSION	GI:4538104			
KEYWORDS	porcine endogenous retrovirus.			
SOURCE	porcine endogenous retrovirus			
ORGANISM	Viruses; Retroid viruses; Retroviridae; Mammalian type C			
	retroviruses; 1-Mammalian type C virus group.			
REFERENCE	1 (bases 1 to 3320)			
AUTHORS	Galbraith,D.N., Haworth,C., Lees,G.M. and Smith,K.T.			
TITLE	PORCINE RETROVIRUS			
JOURNAL	Patent: WO 9740167-A 1 30-OCT-1997;			
	Q ONE BIOTECH LTD (GB)			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

BASE COUNT 922 a 844 c 832 g 722 t
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 3320;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gatggctctctgccccttg 20
|||||
Db 2289 GATGGCTCTCTGCCCCCTTG 2308

RESULT 2

AF038601 7333 bp mRNA MAM 21-MAY-1998
LOCUS Sus scrofa porcine endogenous retrovirus ERV-PK15 mRNA, complete
DEFINITION sequence.
ACCESSION AF038601
VERSION AF038601.1 GI:3133304
KEYWORDS
SOURCE pig.
ORGANISM Sus scrofa

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 7333)
AUTHORS Akiyoshi,D.E., Denaro,M., Zhu,H., Greenstein,J.L., Banerjee,P. and Fishman,J.A.

TITLE Identification of a full-length cDNA for an endogenous retrovirus
of miniature swine

J. Virol. 72 (5), 4503-4507 (1998)

REFERENCE 98216827

2 (bases 1 to 7333)

Fishman,J.A.

TITLE Direct Submission

Submitted (15-DEC-1997) Infectious Disease Unit, Massachusetts
General Hospital, 55 Fruit Street, Boston, MA 02114, USA

REFERENCE 3 (bases 1 to 7333)

Fishman,J.A.

TITLE Direct Submission

Submitted (15-MAY-1998) Infectious Disease Unit, Massachusetts
General Hospital, 55 Fruit Street, Boston, MA 02114, USA

Sequence update by submitter

On May 16, 1998 this sequence version replaced gi:3116445.

REMARK

COMMENT

FEATURES

source

1. 7333

/organism="Sus scrofa"

/strain="miniature swine"

/db_xref="taxon:9823"

/cell_line="PK-15"

/note="porcine endogenous retrovirus ERV-PK15; infective

virus but represents a naturally occurring transcript"

1. 578

598..2172

/codon_start=1

/product="gag protein"

/protein_id="AAC16766.1"

/db_xref="GI:3116446"

/translation="MGQTVTPLSLTLDHWTEVRSRAHLSVQVKKGPWQTFCASEWP

TFDVGWPSGETFSEIILAVKALIFQPGSHDQEPYILTWDLAEDPPVWKPWLN

KPKPGPRLIALGKKNHSAEKVPSPIYPTETPTWPEPPVPPYPAQGVRG

PSAPGAPVAGPAAGTSSRGATPERTDEIALLPTTYGPPMPGQLOPQWPESS

ADLYNKNTHNPPESEDPPQLTGLVESLMSHQPTWDDCOQLQTLPTTERERILLEA

KKNVPGADGRTQIQNEIDMGFPPLTRPGWDYNTAEGRESLKIIYRQALVAGLRGARRP

TNLAKRVNMQGNEPPSPVFLERMEAFRRFTPDPTSEAQKASALAFIGSALDIR

KLQRLGQEAERDLVREAKEYVYRRETEBEQKEKEEREERDRDRQENLT

KILAAVVEGSSRRERDRPKIRSGPRQSGNLNRTPLDKQCAVCKEKGHWARNCPK

KGNKGPVLALEEDKD"

2320..3552

/note="truncated protein"

/codon_start=1

/product="pol protein"

/protein_id="AAC16767.1"

/db_xref="GI:3133305"

/translation="MGATGQROYPWTRTRTVDLGVGRVTHSFLVTPCEPVPVLLGRDLL
TKMQAQSIFEQGRPEVSVNNKPITVLTQLDDEYRLYSQVKKPDODIQSWLEQFPQAW
AETAMGLAKQVPPQVQIQKASATPSVSRQYPLSREAREGIWPHVRIQQGILVQVQ
SPWNTPLLPVRKPGTNDYRPVQDLREVNRQVDIHPTVPNPYLLSALPPERNWTVL
DLKDAFFCLRLHPTSQPLTFEWRDPGTGRGOLTWRLPGQFKNSPTIFDEALHRDL
NFRTHQFQVLLQYVDDLLAGATKQDCLEGTAKALLELELDLGYRASAKKAQICRRP
VNYLGSURGGQRLTEARKKTVVQIPAPTTAKQVREFLSTAGFCRLWIPGFATLAAP
LYPLTKRGGCLPQGGK"

CDS

4803..6725

/note="env protein; N terminus deleted"

/pseudo

/codon_start=1

6726..7333

LTR 1984 a 1788 c 1885 g 1676 t

ORIGIN

Query Match 100.0%; Score 20; DB 4; Length 7333;

Best Local Similarity 100.0%; Pred. No. 3.7;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gatggctctctgccccttg 20

|||||

Db 4442 GATGGCTCTCTGCCCCCTTG 4461

RESULT 3

ARI30474

LOCUS ARI30474 7333 bp DNA PAT 16-MAY-2001

DEFINITION Sequence 2 from patent US 6190861.

ACCESSION ARI30474

VERSION ARI30474.1 GI:14118799

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 7333)

AUTHORS Fishman,J.A.

TITLE Molecular sequences of swine retroviruses method of using

JOURNAL Patent: US 6190861-A 2 20-FEB-2001;

FEATURES

Location/Qualifiers

source

1..7333

/organism="unknown"

BASE COUNT 1984 a 1788 c 1885 g 1676 t

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 7333;

Best Local Similarity 100.0%; Pred. No. 3.7;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gatggctctctgccccttg 20

|||||

Db 4442 GATGGCTCTCTGCCCCCTTG 4461

RESULT 4

A66552

LOCUS A66552 8196 bp DNA PAT 29-MAR-1999

DEFINITION Sequence 2 from Patent WO9740167.

ACCESSION A66552

VERSION A66552.1 GI:4538105

KEYWORDS

SOURCE porcine endogenous retrovirus.

ORGANISM porcine endogenous retrovirus

REFERENCE 1 (bases 1 to 8196)

AUTHORS Galbraith,D.N., Haworth,C., Lees,G.M. and Smith,K.T.

TITLE PORCINE RETROVIRUS

JOURNAL Patent: WO 9740167-A 2 30-OCT-1997;

Q ONE BIOTECH LTD (GB)

FEATURES

Location/Qualifiers

1. .8196
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
BASE COUNT 2165 a 2061 c 2146 g 1820 t 4 others
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 8196;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gatggtctctgccttgg 20
|||||
Db 5226 GATGGCTCTGCGCCTTG 5245

RESULT 5

A66553 8209 bp DNA PAT 29-MAR-1999
DEFINITION Sequence 3 from Patent WO9740167.
ACCESSION A66553
VERSION A66553.1 GI:4538106

porcine endogenous retrovirus.

ORGANISM
Viruses; Retroviral viruses; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.

1 (bases 1 to 8209)

Galbraith, D.N., Haworth, C., Lees, G.M. and Smith, K.T.

PORKINE RETROVIRUS

Patent: WO 9740167-A 3 30-OCT-1997;

Q ONE BIOTECH LTD (GB)

FEATURES

Location/Qualifiers

1. .8209
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
588. .2162
/codon_start=1
/product="VIRAL CORE PROTEIN"
/protein_id="CAB39347.1"
/db_xref="GI:4538107"

/translation="MGQITVPLSLTDHWTVEVSRRAHNSVQVKKGPWQTFCASEWP
TFDVGWSEGTSEIILAVKAIIFOTGSGHPDQEPYILTWQDLADDPWVKKPMWLN
KPRPGRIALGPKNKHSAEKVEPSSYLPDRGADLAGTPTCSPTPLSSSTGCCGEG
TSAPGAPVVEGPAAGTSRRGATPRTDIALPLTYGPPPGGLOPLQWTPSS
ADLNWKNTPPSEDPRLTGVLVESLWFSHQPTWDCQQLQTLFTTERERILLLEA
RKNVPGADGPTQLQNEIDMGFPLRPGWDYNTAEGRESLIYQALVAGLGRASRRP
TNLAKREVWQGPNEPPSVFLMEAFRRFTPDPTSEAQKASVALAFICQALDIR
KKLQRLQGLQAEALDLVREAKEYVYRRETEEEKEREREERERRRROEKLT
KILAAYVEGKSSRRERDRFKIRSGPRQSGNLGNRTPLDKDQAYCKEKGHWARNCPK
GKNGKPVIALDEKD"
/codon_start=1
/product="REVERSE TRANSCRIPTASE"
/protein_id="CAB39348.1"
/db_xref="GI:4538108"

/translation="GRKSDPLPEPRVTLKVEQPFVLDVTGAHSLVLLQPLGKLKE
KSWNGATGQRPWTTRTVDLGVGRVTHSLVTECPVLLGRDLTKMAQISF
EQGPVYVNNKPTITLQDDDEYRLSPQVKPDQDQISWLEQFPQAWAETAGMGLA
KQRPQVYLKASATPVSRQYPLSREAREGIPWVORLIQOGLVPSQPMWTPLLP
VRKPTNDYRPVODLREVNKVDIHTPVNPNYLLSALPPERNWYTLDKDAFTCL
RLHPTSDPLFAFWDPDGTGRTGTLWRLPQGFKNSTIFDEALHRLDANFLRHQHPQ
VTLLQYVDDLLAGATQDCLETKALLLELSDLGHASAKKACICRREYVILGYSLR
GGQWMLTARKKTVQIPAPTAQVREFGTAGFCRLWIPGFATLAAPLYPLTKEG
GFSWAPHOAFADA IKKALLSAPALAPDVTKPTLYVDERKGVARGVLTQTLGWR
PVAYLSKLDLPVAGSWPCLKATAAVALVKDADKLTGONITVIAPHALENTVROP
DRWMTNARMTHYQSLLTERTVAPPALNAPATLPEETDEPVTNHOCHLIEETGVR
KDLTDIPLTEGLTWFTDGSYVVEGKRMAGAAVVDGTRTIWASSLPPEGTSAKAELM
ALTQALRIAGKSNINITYDSRIAFATAHVGAIVKORGLITSAGREIKNEKEITLSLE
ALHPLKRLATIHCPGHQKADLISRGQMADRAKAAQAVNLLPIETPKAPEPRRQ
YTLDEWQETIKIDQFSTPEGTCTVSYTKEILPKHGLEVYVQOIHRLHGLTKHLQOL
VRTSPYHVLRLPGVADSVVVKHCVPCQLVNPANPSRIPPGKRLRGSHPGAHWEVDTEVK

CDS

PARYGNKYLIVFVDFSGWVEAYPTKKETSTVAVAKIILEEIPFPGIPKIVGSDNGPA
FVAQVSOGLAKILGIDWKHCAYRPOSSGOVERMNTIKETLTKLTITGINDMALL
PVLFRVNTPGOGITPYELLYGGPPPIAEAFASADVLLSOPLESRLKALEWVRO
RAWKQLEAYSGDDLOVPHRFQVGSVYVRRHAGNLETRWKPVLVLTPTTAYKVE
GIPTWIIASHVAPPPDPSGWRAEKTEPNLKLRLHLRPLYSNNNSFGQ"
5620. .7590
/codon_start=1
/product="VIRAL ENVELOPE PROTEIN"
/protein_id="CAB39349.1"
/db_xref="GI:4538109"
/translation="MHPTLSRHLPTRGGEKRLRIPLISFASIANFLILITTPQASSK
RLIDSNPHRPLSLTWLIIIDPDTGVTVNSTRGVAPRGTWPELHFLCLRLINPAVKSTP
PNLVSIFYGFCFTEKERYCSCGSEFCRWSVCVTISNDGDKWKLISQDRKVFSEVN
SGPGYKMMKLYKDKSCSPSDLYLKISFTEKTKYKSKVDKWEIGNSFLYLGAG
STLTIRLETGTGEPVAMPKDLAEQGPALPEPHNLVPQLTSLRDPITPPSPNS
TTGLIPTNTPRNSPGVAVTGORLESLOGAFOAINSTDPTDSSCLSSGPPYE
GNAKERKFNVTKEHRNQCWTGSRNKLTLEVSGKTCIGKAPPSHQHLCISYTVVEQA
SENQYLVGNRWACNTGLTFCVSTSVFNQKDFCVWQIVRVYVYHPEEVVLEIDYD
YRNPKEPVSLLTAVMLGLGTAVGVTGTAALITGPQOLKEGLGELHAAMTEDLRA
LKESVNLSESLTSLSEVVLQNRRLGLDLFLREGGLCAALKECCFYVDHSGAIRDMS
NKLKKLERRREREADQGFEGFNRSPWMTLLSALTGPLVLLVLLLTGVPCLINR
FVAFVPERVSAVQIMVLVROQYQGLLSQGETDL"
BASE COUNT 2168 a 2064 c 2154 g 1823 t
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 8209;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gatggtctctgccttgg 20
|||||
Db 5240 GATGGCTCTGCGCCTTG 5259

RESULT 6

PEN133817 8849 bp DNA VRL 07-JAN-2000
LOCUS Porcine endogenous retrovirus type C proviral gag, pol and env
genes and LTR (class A, clone 42).
DEFINITION
ACCESSION AJ133817.1 GI:6688947
VERSION
KEYWORDS env gene; env protein; gag gene; gag protein; pol gene; pol
protein.
SOURCE
ORGANISM
porcine endogenous retrovirus.
Viruses; Retroviral viruses; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
REFERENCE 1 (bases 1 to 8849)
AUTHORS Czauderna, F., Fischer, N., Boller, K., Krach, U., Kurth, R. and
Toenjes, R.R.
TITLE Molecular Characterization of Human-tropic and
Replication-competent Porcine Endogenous Retroviruses
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 8849)
AUTHORS Toenjes, R.R.
TITLE Direct Submission
JOURNAL Submitted (04-MAY-1999) Toenjes R.R., Medical Biotechnology,
Paul-Ehrlich-Institut, Paul-Ehrlich-Strasse 51-59, Hessen, D-63225
Langen, GERMANY
FEATURES
source 1. .8849
/organism="porcine endogenous retrovirus"
/proviral
/strain="type C"
/db_xref="taxon:61673"
/clone="42"
/note="grown in PK15 (ATCC No. CCL-33) cell line"
LTR 1. .668
misc_feature 506
primer_bind 671. .688
/note="gly4"

gene	1115..2686	Matches	20;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
mat_peptide	/gene="gag"	QY	1	gatgctctctgcccctttg	20						
	/gene="gag"	Db	5767	GATGGCTCTCTGCCCCCTTTG	5786						
CDS	/function="retroviral capsid"										
	/product="gag protein"										
	join(1115..2686,2690..6271,6150..8132)										
	/note="viral polypeptide"										
	/codon_start=1										
	/codon_id="CAB65340.1"										
	/db_xref="GI:6688948"										
	/translation="MGQTVTTPLSLTLDHWTVEVRSRAHNLISVQVKKGPWQTCASEWP										
	TFDVGWSEGTENSEIILAVKAIIFOTGCSHPDOEPYILTWODLAEDPPVKKWHLN										
	KPRPGPRILALGCKNKHSAEKVPSPHIYPIELEPPANPEQSVPPPPYDQAARG										
	PSAPGAPVGEAGTARRRGAIPERTDEIALLPRTYPPMPGQLOLYWPFSS										
	ADLYNWKTNHPPFSDPQRLTGLVSLFMSHQPTDDCCQLLTFTTEERKRILLEA										
	RKNVPGADGRPTQLQNEIDMGFPLTRPGWDYNTAEGRESLKIVYKALVAGLRASRRP										
	TNLAKVREVMQGNPEPSPVFLERLMEAFRRPTPDPTSEAKASALAFIGOSALDIR										
	KILQRLGLEQAEELRDVLREAEKVYRRETEEEKOEKEREEREERDRQENKLT										
	KILAAVVECKSSRRERDRPKIRSGPRQSGNLGNRTPLDKQACVCKEKGHWARNCPK										
	KGNKPKVLALEEDKGRGSDPLPEPRVTLKVEGQPVFELVDTAHSHVLLLOPLGKL										
	KEKSSVMGATGQRQYPTWTRTDVLDGVRVTHSVFLVPECPVLLGRDLLKMAQI										
	SFOGPEVSNKPKITVLTQLDDREYRLSPQVKPDQIQSWLEQFQWAEATAGM										
	LKQVPPQVQLKASATPVSVROYPLSREARGLWPHVORLIQOGLVLPVQSPWNTPL										
	LPVRKPTNDYRPVDLREVNKRVDIHTVPNPYNLLSALPERNWTVDLKDAPF										
	CLRLHPTSOPLFAEWDRPTGRTGQLTWRLPQGFKNSTPIDEALHRLDLANFKIOH										
	PQVTLQYVDDLLAGATKQDCGLTKALLLELSLDIGYRASAKKQICRREVTYLGYS										
	LRGGQWLTEARKTVQIAPTAKQVREFLTAGFCRLMIPGFATLAAPLYLTKE										
	KGFSAWPEHQAFDAIKKALLSAPALPDVPTKFTLYVDERKGVARGLVQTILGPW										
	RRPVAYLSKKLDVQSGWPCVKAIAVAAILVKDADKLTGQNTIVIAPHALLIVRO										
	PPDWMTNARMTHYQSLLTERTVYFAPPAALNAPATLPEETDEPETHOCHOLLEETG										
	VRKDLTDPLTCEVLTWFDGSSVVEGKRMAGAAVVDGTRTIWASSLPFGTSAKAE										
	LMALTOALAEKGSINIYDTSKYAFATVHVGAIYKQKGLTSAGRETKKEEILSL										
	LEALHPKRLAIHCPGHQAKDLISRGNMADRAVAKAAQAVNLLPIETPKAPEPR										
	ROYTDEWDEIKIDQFSETPECTCYTSYKEILPHKEGLEYVQIHRHLGLTKHLQ										
	QVTRTSYHVLRLPGVADSVKCVQVNPANPSRIPGKRLRSGHPGAHVEVDTE										
	LPVAKYGNKYLIVVDTFSGWVEAYPTKKTSTVVAKKILEIFRFGIPKVGSDNG										
	PAPVAQVSOGLAKILGIDMKLHCAYRQSSGOVERMNTIKETLLKLTETGINDMA										
	LPFVLFRVNTPGOGGLTPYELLXGGPPPLVEIASVHSADVLLSQPLSRLLALEWV										
	RORAWQLREANSSEGEDLOVPHRFQVDSVYVRRHRAGNLETRWKGPLYLVLLTPTAV										
	KVERITPIHSHVPPAPPPDSGKAEKTENPKLRLHRVVPYSVNNSSMHPTLSRR										
	HPIRGKTKPKRLKIPLSFIAMFLTSLITPQVNGKRLVDSPNSHKPLSLTWLLTDSG										
	YCGNPDDFCQKWCISITSDNGKWPVQODRVSYSFVNNPTSYNOFNYGHRWKDWO										
	QRYOKDVRNKQIJSCHSLDLYLKISFTEKGQENIKQWNGISWIGIYVYGGSRKGS										
	VLTIRLRIETMEPPVAGNGLAEQGPQIQEPRSPNSPDINTSGSVPEPNITI										
	KYCALFLSLIQAFQALNLTTPATSSWCLASGPPYEGMARGKFNVTKEHRDQC										
	TWGSQNKLTLEVSGKGTICIGVPPSSHQHLNHTFAFNRTSESYLVPGYDRWACNT										
	GLTPCVSLTVFNOTKDFCVWQIVPVRVYVPEKAVLDKYDYRNPKRKPEISLTLAVM										
	LGLVAAAGVTGTAALITGPQOLEKLSNLHRTVTEDLQALEKSYNLEESIYLSLEV										
	VIGNRRGLDLLFLKGGGLCVALKKECCFVVDHSGAIRDSMKLRLRRLRRRERADQ										
	GNFEGWNSPMTWTLTLLSALTGPLVLLLLITVGPCLINRFVAFVRVRSVAVQIMVLIR										
	QQQGLLSQGETDL"										
mat_peptide	2690..6271										
	/gene="pol"										
	/function="retroviral protease and polymerase"										
	/product="Pol protein"										
gene	2690..6271										
mat_peptide	/gene="pol"										
	6150..8129										
	/gene="env"										
	/function="retroviral envelope"										
	/product="Env protein"										
gene	6150..8129										
	/gene="env"										
LTR	8182..8849										
polyA_signal	8742..8747										
BASE COUNT	2355 a 2191 c 2281 g 2022 t										
ORIGIN											
Query Match	100.0%;	Score	20;	DB	14;	Length	8849;				
Best Local Similarity	100.0%;	Pred.	No.	3.7;							

Query Match

Best Local Similarity

/translation="GRRGSDPLPEPVRVLKVEGQVPELVDTGAKHSVLLQPLGKLKD
KSKWNGATGQOYQWTRTVDLVGRVTHSFLVPIPECPAPLLGQRLDKTKMAQISF
EKGKPEVSANNKPTVLTILQDEYRLVSPVLPQDNQIFWLEQPPQAWAETAGMGLA
KQVPOVIOKASAAVSVROVPLSKAREGRIRPHVORLQOGLVLPVSPWNTPLLP
VRKPGTNDVRVQDILREVNKRVQDILHPTVPNPNLLCALPORSWYTVLDLKDAPFCL
RLHPTSQPLFAPEWRDPGAGRGQLTWRLPQGFNSPTIDFADLHRLDANFRIOHQ
VILQQLVDLLAGATKQDCLEGTAKALLELSDLGRASAKAICREVTYLYGSLR
TQQLVTEARKETVVOIPAPTTAKQVREFLAGFCRLWIFGFATLAAPLYLTKRG
EFSWAPHOAFDAIKKALLSAPALADPVTKPFTLYVDERKVGAVGLTQTLGPRR
PVALSKLDPVASGWPICAKAIAAVAILVDKDKLTLGONITVIAHALENIVQWRP
DWMNTNARHTVQSLLLTERVTFAALNAPATLLPEETDEPVDHCHOLLIETGVR
KDLTDLPTGEMLTWFTDSSYVWEKRNAGAAVVDGRTIWASSLPEGTSAQKALM
ALTDLPALAEAGSINIYDSTRFAFAHVGAIYKORGLLTSAGREIKKEEISLLE
AHLKPKRAIHCPCGHAKADPISRGNMADRAKAAQVGNLPMIETPKAPEGRQ
YTLDEQELIKIDQFSETPECTYSDGKEILPHKEGLEYYVQIHLRHLGTHLQLOL
VRTSPVHLRLPCVADSVVHCPCOLNANSPRPPGKRLRSGHSPGAHWEVDTEVK
PAKYNKYLVLVFTDSGWEAYPTKSTSTVAAKILIEEIPFREGIPKVGSDNGPA
FYAQSOGKAKILGIDKWLHCAYRPOSSGOVERMRTIKETILKLTETGINDWALL
FVLFRVNTPOQGLTPELILYGGPPLAEIAFAHSAADVLSSQPLFSRIKALEVRQ
RAWKOLREAYSGDLQVPHRFQGVDSVYVRHRAGNLETWKGPYLVLLTPTAVKVE
GIPTWIIASHVAPAPDPGSGWKAETENPLKRLHRVVPVSYVNNSSS"
6185. .8149
/gene="env"
6185. .8149
/gene="env"
/codon_start=1
/product="envelope"
/protein_id="CAC39618.1"
/db_xref="GI:14275840"
/translation="MHPTLSRRHLPIRGKPKRUKIPLSFASIAWFLTSTTPQVNGK
RLVDSPNHKLPLSLTWLLTDSGTGININSTGEAPLGTPWPELYLRSVTPGLNDQA
TPDDVLRAYGFVYCPGNNEEYCGNPDFFCKQWSCITSDNGNMKWPVSQODRVSYS
FYNNPTSVNOFNYGHRKDWORVKDNRKQISCHSDLDLDYLIKISFTEKQENIO
KVNQISGIVYGGSGRKKSVLIRLITOMEPVVAIGPNGLAGQGPPIQDQRP
SNPNDYNTSGSVTEPNITIKGAKLSLIQGAFAALNSTTPEATSSCWLCASGP
PYEGMAGGKFPNVTKEHRDQCTWGSQNKLTITEVSGKGTGCMVPPSHOHLNHTEA
FNRTSEYLVPGYDRWACNTGLTPCYSTLVFNQTKDFCMVQIVPRVYVPEKAVL
DYOALEKSVNLEESLTSLEVLQNRGLDLDLFLKEGLGVALKECCFVVDHSGAI
RDSMKLRELEKHKKEAGQGFEGWFKNSPWPTVLLSALTGPLVLLILLIVGPG
LLNRVAVYREQSVAVRWLVLRQYQGLPS"
8216. .8918
polyA_signal
8810. .8815
BASE COUNT 2435 a 2174 c 2253 g 2056 t

gene

CDS

FEATURES
source

Location/Qualifiers
1. .2687
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:3709119"
/tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
ductal carcinoma, 5 month old virgin mouse."
/clone_lib="NCI_CGAP_Mam6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. .2122
/codon_start=2
/product="Unknown (protein for IMAGE:3709119)"
/protein_id="AAH05647.1"
/db_xref="GI:13542909"

CDS

translation="APVSVPCIPPHRLSKITNLTITQKSFQFMORLHSVTLKRSR
NGVPLRLTHLQSOBCEQVGRDSDKWKALKEOLKQWRHLRDLERALLVELIR
KREKLRETIKIQIATNEMQLTPELLRLKRLQEQKDTGNFSPVPLESEVPDILD
HIRKPMDFMTKONLEAYRYLNFDDFEEDFNVLVSNCLYNAKDTIFYRAVRLRBOG
GAVLQARQARQAKMGIDFETGMHPIHNLADGEVSHHTEDEVEERLLENKHLVPEE
QLKLLERLDEVNASQSVGRSRRAKMKIKETALRKLAHRETGRDGPGRHPSGR
GNLTPHACDKDQOTDAAESSOFTSKGLGNMSSTPAHEVGRRTSVLFKKNPK
TAGPKRPRGPKNRESQMTSPHGGSPVGPOLPMGSLRQKRGSRPSSSSSDSDS
KTSTEDPMPLPANGSSGNQPKKFLYVRCNLPSSSDSESSSSSSSSSSSSAASRT
DTPSGKRGKPSFSGTFPDSSETSENEAYSVGTGRGVGSHVMSVRKSLGRGAGW
LSEDESDPLDALDLVWAKRCGYSPALIIDPKMPREGMHGVPVPPVPLEYKILGE
QMTQEARHELYLVLFDFNKRKTQWLPRTKLVPLGVNQDLDKEKMLBGRKSNIRKSVQI
AYHRLQHRSKVQESSETSDSD"
BASE COUNT 717 a 733 c 740 g 497 t

Query Match 100.0%; Score 20; DB 14; Length 8918;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gatggctctcctgccccttg 20
|||||
Db 5805 GATGGCTCTCCTGCCCCCTTG 5824

RESULT 8

BC005647 2687 bp mRNA ROD 12-JUL-2001
LOCUS Mus musculus, clone IMAGE:3709119, mRNA, partial cds.
DEFINITION BC005647
ACCESSION BC005647
VERSION BC005647.1 GI:13542908
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2687)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (02-APR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

REMARK
COMMENT

USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: villalona@bcm.tmc.edu.
Villalona, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov
Series: IRAP Plate: 17 Row: 1 Column: 9
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.

Location/Qualifiers
1. .2687
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:3709119"
/tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
ductal carcinoma, 5 month old virgin mouse."
/clone_lib="NCI_CGAP_Mam6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. .2122
/codon_start=2
/product="Unknown (protein for IMAGE:3709119)"
/protein_id="AAH05647.1"
/db_xref="GI:13542909"

CDS

translation="APVSVPCIPPHRLSKITNLTITQKSFQFMORLHSVTLKRSR
NGVPLRLTHLQSOBCEQVGRDSDKWKALKEOLKQWRHLRDLERALLVELIR
KREKLRETIKIQIATNEMQLTPELLRLKRLQEQKDTGNFSPVPLESEVPDILD
HIRKPMDFMTKONLEAYRYLNFDDFEEDFNVLVSNCLYNAKDTIFYRAVRLRBOG
GAVLQARQARQAKMGIDFETGMHPIHNLADGEVSHHTEDEVEERLLENKHLVPEE
QLKLLERLDEVNASQSVGRSRRAKMKIKETALRKLAHRETGRDGPGRHPSGR
GNLTPHACDKDQOTDAAESSOFTSKGLGNMSSTPAHEVGRRTSVLFKKNPK
TAGPKRPRGPKNRESQMTSPHGGSPVGPOLPMGSLRQKRGSRPSSSSSDSDS
KTSTEDPMPLPANGSSGNQPKKFLYVRCNLPSSSDSESSSSSSSSSSSSAASRT
DTPSGKRGKPSFSGTFPDSSETSENEAYSVGTGRGVGSHVMSVRKSLGRGAGW
LSEDESDPLDALDLVWAKRCGYSPALIIDPKMPREGMHGVPVPPVPLEYKILGE
QMTQEARHELYLVLFDFNKRKTQWLPRTKLVPLGVNQDLDKEKMLBGRKSNIRKSVQI
AYHRLQHRSKVQESSETSDSD"
BASE COUNT 717 a 733 c 740 g 497 t

Query Match 92.0%; Score 18.4; DB 10; Length 2687;
Best Local Similarity 95.0%; Pred. No. 26;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 gatggctctcctgccccttg 20
|||||
Db 2148 GATGGCTCTCCTGCCCCCTTG 2167

RESULT 9

AX002804 3482 bp DNA PAT 21-AUG-2000
LOCUS AX002804
DEFINITION Sequence 3 from Patent WO9853104.
ACCESSION AX002804
VERSION AX002804.1 GI:9885132
KEYWORDS porcine endogenous retrovirus.
SOURCE porcine endogenous retrovirus
ORGANISM Viruses; Retroviridae; Mammalian type C

```

retroviruses; 1-Mammalian type C virus group.
1 (bases 1 to 3482)
Stoye,J.P. and Weiss,R.A.
Detection of retroviral subtypes based upon envelope specific
sequences
JOURNAL Patent: WO 9853104-A 3 26-NOV-1998;
MEDICAL RES COUNCIL (GB); STOYE JONATHAN PAUL (GB)
FEATURES
    source
        Location/Qualifiers
            /organism="porcine endogenous retrovirus"
            /db_xref="taxon:61673"
BASE COUNT 927 a 854 c 867 g 834 t
ORIGIN
    Query Match 92.0%; Score 18.4; DB 6; Length 3482;
    Best Local Similarity 95.0%; Pred. No. 27;
    Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gatggctctcctgccttgg 20
    ||| ||||| ||||| |||||
Db 528 GATAGCTCTCCTGCCCTTTG 547

RESULT 10
PERENV2 3482 bp RNA 14-OCT-1997
DEFINITION Porcine endogenous retrovirus env gene, 3482 bp.
ACCESSION Y12239
VERSION Y12239.1 GI:2576324
KEYWORDS env gene.
SOURCE porcine endogenous retrovirus.
ORGANISM porcine endogenous retrovirus
    Viruses; Retroviral viruses; Retroviridae; Mammalian type C
    retroviruses; 1-Mammalian type C virus group.
1 (bases 1 to 3482)
Le Tissier,P., Stoye,J.P., Takeuchi,Y., Patience,C. and Weiss,R.A.
Two sets of human-tropic pig retrovirus
JOURNAL Nature 389 (6652), 681-682 (1997)
MEDLINE 97478526
REFERENCE 2 (bases 1 to 3482)
AUTHORS Stoye,J.P.
TITLE Direct Submission
JOURNAL Submitted (02-APR-1997) J.P. Stoye, National Institute for Medical
Research, Virology, The Ridgeway Mill Hill, London, NW7 1AA, UK
REMARK revised by author 16-SEP-97
FEATURES
    source
        Location/Qualifiers
            1..3482
            /organism="porcine endogenous retrovirus"
            /specific_host="Sus scrofa"
            /db_xref="taxon:61673"
            /note="PERV-B"
            911..2884
            /gene="env"
            911..2884
            /gene="env"
            /codon_start=1
            /protein_id="CAA72928.1"
            /db_xref="GI:2576325"
            /db_xref="SPTREMBL:O41173"
            /translation="MHPTLSWRHLPTRGGEKRLRPLSPASIAFWLTLTITPQASSK
            RLIDSSNPHRPLSLTLLIIDPTGVTNTRGVAPRGTWMPHLFCLRLINPAVKSTP
            PNLVRSYGFCCPCTEKYKCGGSGESFCRRSCVTSDGDKWPISLDRVRFSPVN
            SGPKYKVMKLYKDKSCPSLDLYLKISTEKGKQENLOKINMGSKGLVFKYGGGA
            GSTTLRLRTETGTEPPVAVGPKVLAEGQPPALPEPHNLPVQLISLRDITQPNSN
            GTGLIITPRNSPGVPVKTQGLFSLIOGAFAQAINSTDPPDATTSCWCLLSSGPPY
            EGMAKEGVNWKFNWACNTGKTSRKNKLTLEVSKGCTCGKAPPSHQHLCYSTVVEQ
            ASENQYLVPGYNRWACNTGLTPCSTVFNQSKFCVMQIVPRVYVHPHEAVTDEY
            DYRNRPKRREPVSLLTAVMLGIGTAVGVTGTAALITGPOOLEKGLGELHAAMTEDLR
            ALFESVNLSESLTSLSEVLQNRGLDLFLREGGLCAALKKECCFYVDHSGAIRD
            MSKRLRLRRRRRREADQGFEGFNRSFWMNTLSALTGLPLVLLLLLTGVPCLIN
            RFVAFVRRVSVMQIMVLRQOYQGLLSQGETDL"
BASE COUNT 927 a 854 c 867 g 834 t

```

```

ORIGIN
    Query Match 92.0%; Score 18.4; DB 14; Length 3482;
    Best Local Similarity 95.0%; Pred. No. 27;
    Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gatggctctcctgccttgg 20
    ||| ||||| ||||| |||||
Db 528 GATAGCTCTCCTGCCCTTTG 547

RESULT 11
AX052635 4402 bp DNA 12-JAN-2001
LOCUS Sequence 20 from Patent WO0071726.
DEFINITION AX052635
ACCESSION AX052635
VERSION AX052635.1 GI:12226825
KEYWORDS porcine endogenous retrovirus.
SOURCE porcine endogenous retrovirus
    Viruses; Retroviral viruses; Retroviridae; Mammalian type C
    retroviruses; 1-Mammalian type C virus group.
1 (bases 1 to 4402)
Fiederspiel,M.J.
Methods to inhibit infectious agent transmission during
xenotransplantation
JOURNAL Patent: WO 0071726-A 20 30-NOV-2000;
Mayo Medical Ventures (US)
FEATURES
    source
        Location/Qualifiers
            1..4402
            /organism="porcine endogenous retrovirus"
            /db_xref="taxon:61673"
BASE COUNT 1259 a 1085 c 1111 g 947 t
ORIGIN
    Query Match 92.0%; Score 18.4; DB 6; Length 4402;
    Best Local Similarity 95.0%; Pred. No. 27;
    Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gatggctctcctgccttgg 20
    ||| ||||| ||||| |||||
Db 3301 GATAGCTCTCCTGCCCTTTG 3320

RESULT 12
AX052637 4918 bp DNA 12-JAN-2001
LOCUS Sequence 22 from Patent WO0071726.
DEFINITION AX052637
ACCESSION AX052637
VERSION AX052637.1 GI:12226827
KEYWORDS porcine endogenous retrovirus.
SOURCE porcine endogenous retrovirus
    Viruses; Retroviral viruses; Retroviridae; Mammalian type C
    retroviruses; 1-Mammalian type C virus group.
1 (bases 1 to 4918)
Fiederspiel,M.J.
Methods to inhibit infectious agent transmission during
xenotransplantation
JOURNAL Patent: WO 0071726-A 22 30-NOV-2000;
Mayo Medical Ventures (US)
FEATURES
    source
        Location/Qualifiers
            1..4918
            /organism="porcine endogenous retrovirus"
            /db_xref="taxon:61673"
BASE COUNT 1382 a 1173 c 1183 g 1180 t
ORIGIN
    Query Match 92.0%; Score 18.4; DB 6; Length 4918;
    Best Local Similarity 95.0%; Pred. No. 27;

```

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 gatggctctctgcccttg 20
 Db 1804 GATAGCTCTCTGCCCTTG 1823

RESULT 13

AX052636 LOCUS AX052636 6076 bp DNA PAT 12-JAN-2001
 DEFINITION Sequence 21 from Patent WO0071726.
 ACCESSION AX052636
 VERSION AX052636.1 GI:12226826
 KEYWORDS porcine endogenous retrovirus.
 SOURCE porcine endogenous retrovirus.
 ORGANISM Viruses; Retroid viruses; Retroviridae; Mammalian type C
 retroviruses; 1-Mammalian type C virus group.

REFERENCE 1 (bases 1 to 6076)
 AUTHORS Federspiel,M.J.
 TITLE Methods to inhibit infectious agent transmission during
 xenotransplantation
 JOURNAL Patent: WO 0071726-A 21 30-NOV-2000;
 Mayo Medical Ventures (US)

FEATURES
 source
 1..6076
 /organism="porcine endogenous retrovirus"
 /db_xref="taxon:61673"
 BASE COUNT 1613 a 1512 c 1609 g 1342 t
 ORIGIN

Query Match 92.0%; Score 18.4; DB 6; Length 6076;
 Best Local Similarity 95.0%; Pred. No. 28;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 gatggctctctgcccttg 20
 Db 5835 GATAGCTCTCTGCCCTTG 5854

RESULT 14

AX052647 LOCUS AX052647 6076 bp DNA PAT 12-JAN-2001
 DEFINITION Sequence 32 from Patent WO0071726.
 ACCESSION AX052647
 VERSION AX052647.1 GI:12226837
 KEYWORDS porcine endogenous retrovirus.
 SOURCE porcine endogenous retrovirus.
 ORGANISM Viruses; Retroid viruses; Retroviridae; Mammalian type C
 retroviruses; 1-Mammalian type C virus group.

REFERENCE 1 (bases 1 to 6076)
 AUTHORS Federspiel,M.J.
 TITLE Methods to inhibit infectious agent transmission during
 xenotransplantation
 JOURNAL Patent: WO 0071726-A 32 30-NOV-2000;
 Mayo Medical Ventures (US)

FEATURES
 source
 1..6076
 /organism="porcine endogenous retrovirus"
 /db_xref="taxon:61673"
 BASE COUNT 1619 a 1505 c 1598 g 1354 t
 ORIGIN

Query Match 92.0%; Score 18.4; DB 6; Length 6076;
 Best Local Similarity 95.0%; Pred. No. 28;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 gatggctctctgcccttg 20
 Db 5835 GATAGCTCTCTGCCCTTG 5854

RESULT 15

AX052634 LOCUS AX052634 7362 bp DNA PAT 12-JAN-2001
 DEFINITION Sequence 19 from Patent WO0071726.
 ACCESSION AX052634
 VERSION AX052634.1 GI:12226824
 KEYWORDS porcine endogenous retrovirus.
 SOURCE porcine endogenous retrovirus.
 ORGANISM Viruses; Retroid viruses; Retroviridae; Mammalian type C
 retroviruses; 1-Mammalian type C virus group.

REFERENCE 1 (bases 1 to 7362)
 AUTHORS Federspiel,M.J.
 TITLE Methods to inhibit infectious agent transmission during
 xenotransplantation
 JOURNAL Patent: WO 0071726-A 19 30-NOV-2000;
 Mayo Medical Ventures (US)

FEATURES
 source
 1..7362
 /organism="porcine endogenous retrovirus"
 /db_xref="taxon:61673"
 BASE COUNT 1997 a 1821 c 1881 g 1663 t
 ORIGIN

Query Match 92.0%; Score 18.4; DB 6; Length 7362;
 Best Local Similarity 95.0%; Pred. No. 28;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 gatggctctctgcccttg 20
 Db 5900 GATAGCTCTCTGCCCTTG 5919

Search completed: February 24, 2002, 02:56:21
 Job time: 21292 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2002, 03:01:35 ; Search time 584.4 Seconds

(without alignments)
4870.499 Million cell updates/sec

Title: US-09-171-553B-1

Perfect score: 3320

Sequence: 1 gaatcgcggcgcgcgcgcac.....aaaaaaaaaaaaaaaa 3320

Scoring table:

OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 930621 seqs, 428662619 residues

Word size: 0

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

N.Geneseq_1101.*
1: /SIDS2/gcgdata/geneseq/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3320	100.0	3320	18 AAV09698	Porcine retrovirus
2	2682	80.8	8209	18 AAV09700	Porcine retrovirus
3	2666	80.3	8196	18 AAV09699	Porcine retrovirus
4	1091	32.9	7333	22 AAT77726	Detective retrovir
5	1091	32.9	7393	18 AAT74883	Porcine retrovirus
6	848	23.5	4402	22 AAC67020	PERV env protein c
7	797	24.0	6076	22 AAC67021	PERV env protein c
8	797	24.0	7362	22 AAC67019	PERV env protein c
9	786	23.7	2462	20 AAV82748	Pig endogenous ret
10	746	22.5	6076	22 AAC67032	PERV env protein c
11	704	21.2	7892	18 AAT74884	Miniature swine re

12	704	21.2	8132	22 AAT77727	Nucleotide sequenc
13	570	17.2	4918	22 AAC67022	PERV env protein c
14	498	15.0	600	22 AAH26252	Pig endogenous ret
15	497	15.0	599	22 AAH26258	Pig endogenous ret
16	497	15.0	599	22 AAH26260	Pig endogenous ret
17	496	14.9	7873	22 AAH26263	PERV env protein c
18	474	14.3	1980	22 AAC67018	PERV-1-15 env prot
19	466	13.4	599	22 AAH26259	Pig endogenous ret
20	438	13.2	540	22 AAH26253	Pig endogenous ret
21	396	11.9	600	22 AAH26251	Pig endogenous ret
22	396	11.9	600	22 AAH26251	Pig endogenous ret
23	396	11.9	600	22 AAH26257	Pig endogenous ret
24	395	11.9	599	22 AAH26255	Pig endogenous ret
25	395	11.9	599	22 AAH26256	Pig endogenous ret
26	376	11.3	478	22 AAH26263	Pig endogenous ret
27	375	11.3	8060	18 AAT74811	Porcine retrovirus
28	375	11.3	8060	22 AAT77725	Tsukuba-1 CDNA. P
29	356	10.7	458	22 AAH26264	Pig endogenous ret
30	325	9.8	478	22 AAH26265	Pig endogenous ret
31	321	9.7	3482	20 AAV82749	Pig endogenous ret
32	301	9.1	478	22 AAH26266	Pig endogenous ret
33	275	8.3	478	22 AAH26262	Pig endogenous ret
34	229	6.9	458	22 AAH26261	Pig endogenous ret
35	45	1.4	45	22 AAC82276	Human retrovirus D
36	41	1.2	45	22 AAC82275	Human retrovirus D
37	35	1.1	127	22 AAC86002	DOREV RT sequence,
38	35	1.1	8655	20 AAV69750	Mus dunali endogeno
39	32	1.0	463	22 AAH71398	Human cervical can
40	31	0.9	628	20 AA207249	Mouse telomerase R
41	31	0.9	955	22 AAF91898	Human secreted pro
42	31	0.9	955	22 AAF91925	Human secreted pro
43	31	0.9	1010	21 AAT18191	Lung cancer associ
44	31	0.9	1404	7 AAK0283	Sequence encoding
45	31	0.9	1404	8 AAN70505	Sequence encoding

ALIGNMENTS

RESULT 1	
ID	AAV09698 standard; DNA; 3320 BP.
XX	
AC	AAV09698:
XX	
DT	19-MAY-1998 (first entry)
XX	
DE	Porcine retrovirus pol and env DNA.
XX	
KW	Porcine retrovirus: POL; POL; protein; ENV; protein; vaccine;
KW	diagnosis; xenotransplantation; prophylactic; therapeutic; ds.
XX	
OS	Porcine retrovirus.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	23..2793
FT	/*tag= a
FT	/product= POL; protein
FT	/note= "polymerase protein"
FT	2642..3397
FT	/*tag= b
FT	/product= ENV; protein
FT	/note= "envelope protein"
XX	
PN	MO9740167-A1.
XX	
PD	30-OCT-1997.
XX	
PF	18-APR-1997; 97MO-GB01087.
XX	
PR	10-FEB-1997; 97GB-0002668.
PR	19-APR-1996; 96GB-0008164.
XX	

PA (IMUT-) IMUTRAN LTD.
PA (OONE-) Q-ONE BIOTECH LTD.

PI Galbraith DN, Haworth C, Lees GM, Smith KT;

XX WPI; 1997-535851/49.

XX Polynucleotide encoding porcine retrovirus expression product -
PT useful to develop products for use in vaccines, diagnosis and
PT xeno-transplantation

XX Claim 4: Fig 1: 69pp; English.

XX This DNA sequence encodes the porcine retrovirus (PoEV) polymerase (POL)
CC and envelope (ENV) proteins. These proteins can be used to develop viral
CC vaccines, antisense nucleic acids, ribozymes and other antiviral agents.
CC They can also be used in xeno-transplantation technology and as
CC diagnostic tools.

XX Sequence 3320 BP; 922 A; 844 C; 832 G; 722 T; 0 other;

Query Match 100.0%; Score 3320; DB 18; Length 3320;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gaattcgccgagcgtcgacagatgctctctctgctgctgagatcacccactagccaa 60
DB 1 gaattcgccgagcgtcgacagatgctctctctgctgctgagatcacccactagccaa 60
QY 61 ccaatttttgccttcgaatggaagatccaggtacggaagaacccgagcagctcacctg 120
DB 61 ccaatttttgccttcgaatggaagatccaggtacggaagaacccgagcagctcacctg 120
QY 121 acccgactgccccagaggttcaagaactcccccacatctttgacgaagcccttaacagg 180
DB 121 acccgactgccccagaggttcaagaactcccccacatctttgacgaagcccttaacagg 180
QY 181 gacctggccaactcagatccacacccctcagtgacctccctccagtagctgagac 240
DB 181 gacctggccaactcagatccacacccctcagtgacctccctccagtagctgagac 240
QY 241 ctgctcttcgagcagccacccaacagactgctttagaagtagcagaagcactactgctg 300
DB 241 ctgctcttcgagcagccacccaacagactgctttagaagtagcagaagcactactgctg 300
QY 301 gaattgtctgacctagctacagagctctgcttagaagaagcccgacttttcagagagag 360
DB 301 gaattgtctgacctagctacagagctctgcttagaagaagcccgacttttcagagagag 360
QY 361 gtaacatacttgggtatcaggttctggggcgggcgagcgatgctggaagggagccggaag 420
DB 361 gtaacatacttgggtatcaggttctggggcgggcgagcgatgctggaagggagccggaag 420
QY 421 aaactgtagtcagatccgcccacccaacagcagccaaacagtgagagagtttttgggg 480
DB 421 aaactgtagtcagatccgcccacccaacagcagccaaacagtgagagagtttttgggg 480
QY 481 acaagctgattttgacagctgtagatcccgaggttttgacacttagagcccccactac 540
DB 481 acaagctgattttgacagctgtagatcccgaggttttgacacttagagcccccactac 540
QY 541 ccgctaaacaaagaaagggggaattccctgggctcctctgagccacccaagagcatttgat 600
DB 541 ccgctaaacaaagaaagggggaattccctgggctcctctgagccacccaagagcatttgat 600
QY 601 gctalcaaaaagccctctgagcgcaactgctctggccctccctgagcgtactaaacc 660
DB 601 gctalcaaaaagccctctgagcgcaactgctctggccctccctgagcgtactaaacc 660
QY 661 tttaaccttatgtgtagagcgtaaggaatgacccgagagagtttaacccaacctta 720
DB 661 tttaaccttatgtgtagagcgtaaggaatgacccgagagagtttaacccaacctta 720

QY 721 ggacatgagagagacctgttgcctactgtcaaaagagcttgatccctgtagcagtggt 780
DB 721 ggacatgagagagacctgttgcctactgtcaaaagagcttgatccctgtagcagtggt 780
QY 781 tggcccgatgtctgaagagctatcgagctgtgcccatactggttcaagaagcgtgacaaa 840
DB 781 tggcccgatgtctgaagagctatcgagctgtgcccatactggttcaagaagcgtgacaaa 840
QY 841 ttgacttggagcaataataactgtaataagcccccatgcatcttgagaacatcttgcg 900
DB 841 ttgacttggagcaataataactgtaataagcccccatgcatcttgagaacatcttgcg 900
QY 901 cagccccagacagatggaatgacacagccgcagatgacccaatcaaaagcctgtctc 960
DB 901 cagccccagacagatggaatgacacagccgcagatgacccaatcaaaagcctgtctc 960
QY 961 acagagagaggtacttgcctccacagccgctctcaacccctgacactcttctgctgaa 1020
DB 961 acagagagaggtacttgcctccacagccgctctcaacccctgacactcttctgctgaa 1020
QY 1021 ggaactgataaacagatgacatgatgacacacacatgattgagagagctgggtc 1080
DB 1021 ggaactgataaacagatgacatgatgacacacacatgattgagagagctgggtc 1080
QY 1081 cgaagagacctaagac 1140
DB 1081 cgaagagacctaagac 1140
QY 1141 agcagctatgtgtgtaagagtaagagatggtgtggggcgaggtgtggaagggagccgc 1200
DB 1141 agcagctatgtgtgtaagagtaagagatggtgtggggcgaggtgtggaagggagccgc 1200
QY 1201 acgacttggccagcagccttcgcgcggaaggaactcagcgcgaagagctgagctcagc 1260
DB 1201 acgacttggccagcagccttcgcgcggaaggaactcagcgcgaagagctgagctcagc 1260
QY 1261 ctcaagcaagcttgcgcgtgcgcggaaggaatccataacattatccagcagcagc 1320
DB 1261 ctcaagcaagcttgcgcgtgcgcggaaggaatccataacattatccagcagcagc 1320
QY 1321 tatgcttggcagctgacacgttaccggtgacatcattataacaaaggggttgcctacc 1380
DB 1321 tatgcttggcagctgacacgttaccggtgacatcattataacaaaggggttgcctacc 1380
QY 1381 tcagcagagaggaataaagaacaaagaggaattctaaagcctatlaagaagccttaacat 1440
DB 1381 tcagcagagaggaataaagaacaaagaggaattctaaagcctatlaagaagccttaacat 1440
QY 1441 ttgccaanaaggttagctatataacatgctcctgagacatcagaagccaaagatccata 1500
DB 1441 ttgccaanaaggttagctatataacatgctcctgagacatcagaagccaaagatccata 1500
QY 1501 tctaaaggggaacagatgctgtagccggtttgccaagcagagccagagctgttaacct 1560
DB 1501 tctaaaggggaacagatgctgtagccggtttgccaagcagagccagagctgttaacct 1560
QY 1561 ctgctcataatagaacacgcccagaagccagaaacccagacagcagctagagaagc 1620
DB 1561 ctgctcataatagaacacgcccagaagccagaaacccagacagcagctagagaagc 1620
QY 1621 ttgcaagagataaanaagataagacaggttctctgagacccggaaggggagcctgataac 1680
DB 1621 ttgcaagagataaanaagataagacaggttctctgagacccggaaggggagcctgataac 1680
QY 1681 tcatatgggaaggaataatccctgcccacaaagaggttgaataatgctcaacagataat 1740
DB 1681 tcatatgggaaggaataatccctgcccacaaagaggttgaataatgctcaacagataat 1740
QY 1741 cgtctaacccacttagaacataacacctgacagcagttgtgtaagacatcccttataat 1800
DB 1741 cgtctaacccacttagaacataacacctgacagcagttgtgtaagacatcccttataat 1800

QY 1801 gtctgaggtacacagagtggtgactcgggtgtgtcacaacattgtgtccctgccaagt 1860
 Db 1801 gtctgaggtacacagagtggtgactcgggtgtgtcacaacattgtgtccctgccaagt 1860
 QY 1861 gttatgttaattcttccagaatactccacagaaagagactaaagggaaagccaccagcc 1920
 Db 1861 gttatgttaattcttccagaatactccacagaaagagactaaagggaaagccaccagcc 1920
 QY 1921 gctactggaagtggagactcactgaggttaaagccggctaataccggaacaaatatacta 1980
 Db 1921 gctactggaagtggagactcactgaggttaaagccggctaataccggaacaaatatacta 1980
 QY 1981 ttggttttctgtagaacactttccagatggttagaggtctatccctactaaagaaagagact 2040
 Db 1981 ttggttttctgtagaacactttccagatggttagaggtctatccctactaaagaaagagact 2040
 QY 2041 tcaacccttggtggttaagaaatactctgtagagaaattttccaaagatttggataactaaag 2100
 Db 2041 tcaacccttggtggttaagaaatactctgtagagaaattttccaaagatttggataactaaag 2100
 QY 2101 gtaataaggttcaagcaaatgttccagcttctgtgtccaggtaaagtcaaggtcggccaag 2160
 Db 2101 gtaataaggttcaagcaaatgttccagcttctgtgtccaggtaaagtcaaggtcggccaag 2160
 QY 2161 atattggtggtatgtatggaactgtgcatgtgtacacagaccacaaagtctcagagcagta 2220
 Db 2161 atattggtggtatgtatggaactgtgcatgtgtacacagaccacaaagtctcagagcagta 2220
 QY 2221 gaggaggtatgaataaacatataagagagacccttaaccacattgaccacagagactgtgcat 2280
 Db 2221 gaggaggtatgaataaacatataagagagacccttaaccacattgaccacagagactgtgcat 2280
 QY 2281 aatgatgtgagtggtcctcgtccctgtgtctttagaggtgtagaagacacccttggaag 2340
 Db 2281 aatgatgtgagtggtcctcgtccctgtgtctttagaggtgtagaagacacccttggaag 2340
 QY 2341 ttgtggtctgacccctctatgtaattgtctacggtggaagcccccgtgtgagaatgtgc 2400
 Db 2341 ttgtggtctgacccctctatgtaattgtctacggtggaagcccccgtgtgagaatgtgc 2400
 QY 2401 ttgtcacatagtgtgtagtggtcgtcttccacgcttctgtctgtgctgagctcaagggctc 2460
 Db 2401 ttgtcacatagtgtgtagtggtcgtcttccacgcttctgtctgtgctgagctcaagggctc 2460
 QY 2461 gagtgtggtgtagcagcgagcggtggaagcagctccggaaggtcactcagagagagagact 2520
 Db 2461 gagtgtggtgtagcagcgagcggtggaagcagctccggaaggtcactcagagagagagact 2520
 QY 2521 caagttccacatcgtcttccaaagtgtgagatcagctctagtgtagacgcgcagtgacaga 2580
 Db 2521 caagttccacatcgtcttccaaagtgtgagatcagctctagtgtagacgcgcagtgacaga 2580
 QY 2581 aacctcgaagctcgtgtggaagggacctatctgtaacttttggacacacaaaggtgtg 2640
 Db 2581 aacctcgaagctcgtgtggaagggacctatctgtaacttttggacacacaaaggtgtg 2640
 QY 2641 aaagtctgaaggaatcccccacttgatccatgcaatcccaagcttgaagcgcgccactccc 2700
 Db 2641 aaagtctgaaggaatcccccacttgatccatgcaatcccaagcttgaagcgcgccactccc 2700
 QY 2701 gattcgggtgtggaagcgccgaagaaagactgtaaaatcccccctaagtctgcgtccatcgctg 2760
 Db 2701 gattcgggtgtggaagcgccgaagaaagactgtaaaatcccccctaagtctgcgtccatcgctg 2760
 QY 2761 gtctctactctgtcaataactcctcaagttaattgtttaagcgtgtgtagacagccgaa 2820
 Db 2761 gtctctactctgtcaataactcctcaagttaattgtttaagcgtgtgtagacagccgaa 2820
 QY 2821 ctccataaacccttatctctcaactgttactactgactccggaagattataat 2880
 Db 2821 ctccataaacccttatctctcaactgttactactgactccggaagattataat 2880
 QY 2881 taacagcactcaaggggaggtcctcctgtgggagctgtgtgcctgaattatagtcgtcct 2940

Db 2881 taacagcactcaaggggaggtcctcctgtgggagctgtgtgcctgaattatagtcgtcct 2940
 QY 2941 tggataagtaattcctcgtgtctcaatgacagggccacaccccccgatgtactccgtgtcta 3000
 Db 2941 tggataagtaattcctcgtgtctcaatgacagggccacaccccccgatgtactccgtgtcta 3000
 QY 3001 cgggttttcgtgttgcacagggacccccaataatgagaatgtgtggaatccctcaga 3060
 Db 3001 cgggttttcgtgttgcacagggacccccaataatgagaatgtgtggaatccctcaga 3060
 QY 3061 ttctccttgcaagcaatgtagctgtcaatacttcaatgtagtgggaattggaatggccagt 3120
 Db 3061 ttctccttgcaagcaatgtagctgtcaatacttcaatgtagtgggaattggaatggccagt 3120
 QY 3121 ctctcagcaagaagagtagtactcctttgttaacaactcctaccagttataatcaat 3180
 Db 3121 ctctcagcaagaagagtagtactcctttgttaacaactcctaccagttataatcaat 3180
 QY 3181 taattatggccatggtgagatggaagaattgtgcaacagcggtacaaagaatgtacgaa 3240
 Db 3181 taattatggccatggtgagatggaagaattgtgcaacagcggtacaaagaatgtacgaa 3240
 QY 3241 taagcaataaagctgtcattcgtttagacctagattacttaaaataaagtttcaactaaaa 3300
 Db 3241 taagcaataaagctgtcattcgtttagacctagattacttaaaataaagtttcaactaaaa 3300
 QY 3301 aaaaaaaaaaaaaaaaaa 3320
 Db 3301 aaaaaaaaaaaaaaaaaa 3320

RESULT 2
 AAV09700
 ID AAV09700 standard; DNA; 8209 BP.
 XX AAV09700;
 AC AAV09700;
 DT 20-MAY-1998 (first entry)
 XX
 DE Porcine retrovirus DNA encoding the LTR and GAG, POL and ENV proteins.
 XX
 KW Porcine retrovirus; POEV; POL protein; ENV protein; GAG protein;
 KW virion core polypeptide; polymerase protein; envelope protein;
 KW vaccine; diagnosis; xenotransplantation; prophylactic; therapeutic; ds.
 XX
 OS Porcine retrovirus.
 XX
 FH Key
 FT misc_feature
 FT Location/Qualifiers
 FT 1..61
 FT /tag= a
 FT /note= "R LTR domain"
 FT LTR
 FT 1..588
 FT /tag= b
 FT /note= "u5 LTR domain"
 FT misc_feature
 FT 62..143
 FT /tag= c
 FT /note= "u5 LTR domain"
 FT CDS
 FT 588..2162
 FT /tag= d
 FT /product= GAG protein
 FT /note= "viral core polypeptide"
 FT CDS
 FT 2163..5747
 FT /tag= e
 FT /product= POL protein
 FT /note= "polymerase peptide"
 FT CDS
 FT 5748..7590
 FT /tag= f
 FT /product= ENV protein
 FT /note= "envelope protein"
 FT LTR
 FT 7591..8209
 FT /tag= g
 FT /note= "envelope protein"
 FT misc_feature
 FT 8210..8209
 FT /tag= h

FT	protein_bind	/note="U3 LTR domain"
FT		7643..7686
FT		/*tag= i
FT		/note="Binding site for transcription factors
FT	protein_bind	Oct-1, c-Myb, Myf-1, E47 and ETS-1"
FT		7719..7739
FT		/*tag= j
FT		/note="Binding site for transcription factors
FT	protein_bind	AP-4 and NF1"
FT		7744..7754
FT		/*tag= k
FT		/note="Binding site for transcription factor AP-1/TR"
FT	protein_bind	7773..7785
FT		/*tag= l
FT		/note="Binding site for transcription factor ETS-1/GATA"
FT	protein_bind	7793..7834
FT		/*tag= m
FT		/note="Binding site for transcription factors
FT	protein_bind	GATA, ETS-1, c-Myb, AP-1, GATA, E47"
FT		7841..7851
FT		/*tag= n
FT		/note="Binding site for transcription factor AP-1"
FT	protein_bind	7873..7883
FT		/*tag= o
FT		/note="Binding site for transcription factor AP-1"
FT	protein_bind	7885..7914
FT		/*tag= p
FT		/note="Binding site for transcription factors
FT	protein_bind	AP-1/GATA, c-Myb and AP-1"
FT		7928..7953
FT		/*tag= q
FT		/note="Binding site for transcription factors E47
FT	CAAT_signal	7964..7968
FT		/*tag= r
FT	protein_bind	7985..7997
FT		/*tag= s
FT		/note="Binding site for transcription factor GATA"
FT	CAAT_signal	8003..8007
FT		/*tag= t
FT	protein_bind	8024..8036
FT		/*tag= u
FT		/note="Binding site for transcription factor GATA"
FT	CAAT_signal	8042..8046
FT		/*tag= v
FT	protein_bind	8049..8060
FT		/*tag= w
FT		/note="Binding site for transcription factor
FT	protein_bind	AP-1/CREB"
FT		8075..8090
FT	TATA_signal	/*tag= x
FT		1..61
FT		/*tag= y
FT		/note="R LTR domain"
FT	misc_feature	8127..8135
FT		/*tag= z
FT	misc_feature	/note="PolyA downstream element"
FT		8161..8166
FT	polyA_signal	/*tag= aa
XX		
XX	W09740167-A1.	
XX		
XX	30-OCT-1997.	
XX		
XX	18-APR-1997;	97WC-GB01087.
XX		
XX	10-FEB-1997;	97GB-0002668.
XX	19-APR-1996;	96GB-0008164.
XX		
XX	(IMUT-) IMUTRAN LTD.	
XX	(OONE-) O-ONE BIOTECH LTD.	
XX		
XX	Galbraith DN, Haworth C, Lees GM, Smith KT;	

XX	WP1: 1997-535851/49.
DR	P-ESDB: AAM39271; AAM39272; AAM39273.
XX	
XX	Polynucleotide encoding porcine retrovirus expression product -
PT	useful to develop products for use in vaccines, diagnosis and
PT	xeno-transplantation
XX	
PS	Claim 4; Fig 3; 69pp; English.
XX	
CC	This DNA sequence encodes the porcine retrovirus (POEV) virion core
CC	polypeptide (GAC), polymerase (POL) and envelope (ENV) proteins and
CC	also includes the Long Terminal Repeat (LTR). These proteins can be used
CC	to develop viral vaccines, antitense nucleic acids, ribozymes and other
CC	antiviral agents. They can also be used in xeno-transplantation
CC	technology and as diagnostic tools.
XX	
SQ	Sequence 8209 BP; 2168 A; 2064 C; 2154 G; 1823 T; 0 other;
	Query Match 80.8%; Score 2682; DB 18; Length 8209;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 2682; Conservative 0; Mismatches 0; Indels 0; Gaps
OY	21 agatgctctctctgctcgatgattaaacccactatgccaacacttttgcctcgaa
DB	2972 agatgctctctctgctcgatgattaaacccactatgccaacacttttgcctcgaa
OY	81 gagagatccaggtacgggaaagaccgggagcctacacttgaaacccgactgccc
DB	3032 gagagatccaggtacgggaaagaccgggagcctacacttgaaacccgactgccc
OY	141 caagaactcccgacacactctttgacgaagcctacacaggaagacttggccaact
DB	3092 caagaactcccgacacactctttgacgaagcctacacaggaagacttggccaact
OY	201 ccaaacacctaggttgacctctctccagatcagtgatgacctgtcttcggcgag
DB	3152 ccaaacaccttaggttgacctctctccagatcagtgatgacctgtcttcggcgag
OY	261 caaacagacactgcttaaaaggtacgaagagacactacgctggaatttcgact
DB	3212 caaacagacactgcttaaaaggtacgaagagacactacgctggaatttcgact
OY	321 cagagcctctgctaaagaagcccaagattgacgaagagaggttaacatactt
DB	3272 cagagcctctgctaaagaagcccaagattgacgaagagaggttaacatactt
OY	381 tttcggggcgcgagcagatgctgctgaagcgagcaggaagaaactgtatcc
DB	3332 tttcggggcgcgagcagatgctgctgaagcgagcaggaagaaactgtatcc
OY	441 ggccccaacacagccaacaacagltgagagagtttttgggagcagctgga
DB	3392 ggccccaacacacagccaacaacagltgagagagtttttgggagcagctgga
OY	501 gtgagatccgggggttttcgaccttagagagcccaactcaccgcttaacaa
DB	3452 gtgagatccgggggttttcgaccttagagagcccaactcaccgcttaacaa
OY	561 gggattctcctgagctctcttgagcaccgaagcatttgaatgcatataa
DB	3512 gggattctcctgagctctcttgagcaccgaagcatttgaatgcatataa
OY	621 gagcgacactgctcttgacctctccctgaagtaactaaaccccttaccct
DB	3572 gagcgacactgctcttgacctctccctgaagtaactaaaccccttaccct
OY	681 ggcgtaaaggagttacccgagagattttaaaccnaaccccttaggacata
DB	3632 ggcgtaaaggagttacccgagagattttaaaccnaaccccttaggacata
OY	741 tgcctactctgtcaagaagacttgatccctgtagccagtggttggccgtagt

XX	Key	Location/Qualifiers
FH	CDS	576..2126
FT		/*tag= a
FT		/product= GAG protein
FT		/note= "virion core polypeptide"
FT		2143..5733
FT		/*tag= b
FT		/product= POL protein
FT		/note= "polymerase peptide sequence as given in
FT		specification"
FT		5606..7576
FT		/*tag= c
FT		/product= ENV protein
FT		/note= "envelope protein"
XX		
PN		W09740167-A1.
XX		
PD		30-OCT-1997.
XX		
PF		18-APR-1997; 97WO-GB01087.
XX		
PR		10-FEB-1997; 97GB-0002668.
XX		
PR		19-APR-1996; 96GB-0008164.
XX		
PA		(IMUT-) IMUTRAN LTD.
XX		(OONE-) O-ONE BIOTECH LTD.
PI		Galbraith DN, Haworth C, Lees GM, Smith KT;
XX		
DR		WPI; 1997-535851/49.
XX		
PR		Polynucleotide encoding porcine retrovirus expression product -
PT		useful to develop products for use in vaccines, diagnosis and
PT		xeno-transplantation
XX		
PS		Claim 4; Fig 2; 6pp: English.
XX		
CC		This DNA sequence encodes the porcine retrovirus (PoEV) virion core
CC		polypeptide (GAG), polymerase (POL) and envelope (ENV) proteins. These
CC		proteins can be used to develop viral vaccines, antisense nucleic acids,
CC		ribozymes and other antiviral agents. They can also be used in
CC		xeno-transplantation technology and as diagnostic tools.
XX		
SQ		Sequence 8196 BP; 2165 A; 2061 C; 2147 G; 1820 T; 3 other;
Query Match 80.3%; Score 2666; DB 18; Length 8196;		
Best Local Similarity 100.0%; Pred. No. 0;		
Matches 2666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	21	agatgcctctctctgcctgagattacaccccaactgccaacacatttgccttgaatg 80
DB	2958	agatgcctctctctgcctgagattacaccccaactgccaacacatttgccttgaatg 3017
QY	81	ggagatccaggtaacgggaagaccgggagctacactgagaccgactgcccgaaggtt 140
DB	3018	ggagatccaggtaacgggaagaccgggagctacactgagaccgactgcccgaaggtt 3077
QY	141	caagaactcccgacatctttgacgaagccctacacagaggaactgccaactgaagt 200
DB	3078	caagaactcccgacatctttgacgaagccctacacagaggaactgccaactgaagt 3137
QY	201	caaacacccctcaggtaacctctcctcagtaagctgtaagctgcttgcggagaccac 260
DB	3138	caaacacccctcaggtaacctctcctcagtaagctgtaagctgcttgcggagaccac 3197
QY	261	caaacaggaactgttaagaagtaagaagcactactgctggaattgtctgacctaggcta 320
DB	3198	caaacaggaactgttaagaagtaagaagcactactgctggaattgtctgacctaggcta 3257
QY	321	cgagagcctctgtcaaaaagcccgagattgcagggagagggtaaatatttgggtacag 380

Dh	3258	cagagccctctgctaagaagcccaaggtttgcagngagagugfatacatcttggggttaacag	3317
Qy	381	tttcgvggvcgvgcagatgctgcagagcagcaggaagaaactgtatgctccagatacc	440
Dh	3318	tttcgvggvcgvgcagatgctgcagagcagcaggaagaaactgtatgctccagatacc	3377
Qy	441	ggccccaacccaacgaacgaatgtagagagtttttgggagacagcttgatatttgagact	500
Dh	3378	ggccccaacccaacgaacgaatgtagagagtttttgggagacagcttgatatttgagact	3437
Qy	501	gtgatacccgvggttttgagacctttagacagcccatctacccgctataccaagaagaaag	560
Dh	3438	gtgatacccgvggttttgagacctttagacagcccatctacccgctataccaagaagaaag	3497
Qy	561	gggatttctctgvgctctctgaagccaagaagagattttagttatcaaaaagccctct	620
Dh	3498	gggatttctctgvgctctctgaagccaagaagagattttagttatcaaaaagccctct	3557
Qy	621	gagcgcaactgctcttgagccctccctagcgttaactaaaccttataccttlatgtatga	680
Dh	3558	gagcgcaactgctcttgagccctccctagcgttaactaaaccttataccttlatgtatga	3617
Qy	681	ggtataagvgatagcccgagaggttttaacccaacctatgagccaatgtagagagaccgt	740
Dh	3618	ggtataagvgatagcccgagaggttttaacccaacctatgagccaatgtagagagaccgt	3677
Qy	741	tgccctaactgtccaagaagagcttgatccctgtagcagatgtttgcgcgtatgataagc	800
Dh	3678	tgccctaactgtccaagaagagcttgatccctgtagcagatgtttgcgcgtatgataagc	3737
Qy	801	tatcgacagctgttgccaactactgtgtcaagagcgtgcgaacattgacttttgtagcaga	860
Dh	3738	tatcgacagctgttgccaactactgtgtcaagagcgtgcgaacattgacttttgtagcaga	3797
Qy	861	aactgttaatgacccccaatgactatgttagagaaactcgtttgcgacgccccagaccggtat	920
Dh	3798	aactgttaatgacccccaatgactatgttagagaaactcgtttgcgacgccccagaccggtat	3855
Qy	921	gaccacaaccccgcaatgacccaactatcaaaaagcctgtctctcaacagagaggttcaacttcgc	980
Dh	3856	gaccacaaccccgcaatgacccaactatcaaaaagcctgtctctcaacagagaggttcaacttcgc	3917
Qy	981	tccaacacgcgctctaaacctgtgcacacttctgccttgtaagagagactgtatgaacccagtgac	1040
Dh	3918	tccaacacgcgctctaaacctgtgcacacttctgccttgtaagagagactgtatgaacccagtgac	3977
Qy	1041	tcatattgccaatacaactatgtatctgtagagagacttgggtctccgaagaagaccttacaagacat	1100
Dh	3978	tcatattgccaatacaactatgtatctgtagagagacttgggtctccgaagaagaccttacaagacat	4037
Qy	1101	accgcttgacgtgagagatgttcaaccgtgttcaactgacggaagcagctatgtgttggaag	1160
Dh	4038	accgcttgacgtgagagatgttcaaccgtgttcaactgacggaagcagctatgtgttggaag	4097
Qy	1161	taaagagatgtctgtagcagcagtggttggaacggaacccgcagacatcttgtagcagaagcct	1220
Dh	4098	taaagagatgtctgtagcagcagtggttggaacggaacccgcagacatcttgtagcagaagcct	4155
Qy	1221	ggcggaaaggaacttaacgcgaaaaagcgtgagctatagcccttaacgaagctttgcgct	1280
Dh	4158	ggcggaaaggaacttaacgcgaaaaagcgtgagctatagcccttaacgaagctttgcgct	4217
Qy	1281	ggcggaaaggaatccataaacatttatacgaagaagaatgtagcctttgtagactgcaca	1344
Dh	4218	ggcggaaaggaatccataaacatttatacgaagaagaatgtagcctttgtagactgcaca	4277
Qy	1341	cgtaacacggggccatctataacaagaagggtgtgttaccctacgacgagaggaataaa	1400
Dh	4278	cgtaacacggggccatctataacaagaagggtgtgttaccctacgacgagaggaataaa	4337
Qy	1401	gaacaagaaggaatattcaagacctatagagaccttatctttgccaanaagcttaactat	1460
Dh	4338	gaacaagaaggaatattcaagacctatagagaccttatctttgccaanaagcttaactat	4397

QY 1461 tatacactgtcccttgacatcagaaagccaaagatcctatctagtagaggaagccagatgac 1520
 Db 4398 tatacactgtcccttgacatcagaaagccaaagatcctatctagtagaggaagccagatgac 4457
 QY 1521 tgaccgggttgcacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1580
 Db 4458 tgaccgggttgcacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4517
 QY 1581 caaagcccaagaaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1640
 Db 4518 caaagcccaagaaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4577
 QY 1641 agaccagttctctgagactccggaggggagcctgatacctatcctatctagtagaggaagcagc 1700
 Db 4578 agaccagttctctgagactccggaggggagcctgatacctatcctatctagtagaggaagcagc 4637
 QY 1701 gccccacaagaaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1760
 Db 4638 gccccacaagaaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4697
 QY 1761 taacacactgacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1820
 Db 4698 taacacactgacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4757
 QY 1821 ggcctgactcgggtggtcacaacatgtgtgcccctgacagcagcagcagcagcagcagcagc 1880
 Db 4758 ggcctgactcgggtggtcacaacatgtgtgcccctgacagcagcagcagcagcagcagcagc 4817
 QY 1881 aatccctccaggaagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1940
 Db 4818 aatccctccaggaagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4877
 QY 1941 cactgaggtaaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2000
 Db 4878 cactgaggtaaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4937
 QY 2001 ttcaagatggtgtagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2060
 Db 4938 ttcaagatggtgtagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4997
 QY 2061 aatctctgaggaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2120
 Db 4998 aatctctgaggaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5057
 QY 2121 tccagcttcgttgcacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2180
 Db 5058 tccagcttcgttgcacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5117
 QY 2181 actgcatgtgcatatagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2240
 Db 5118 actgcatgtgcatatagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5177
 QY 2241 taaagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2300
 Db 5178 taaagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5237
 QY 2301 gcccttctgtcttctttagaggtgaggaacaccccttgacagcagcagcagcagcagcagcagc 2360
 Db 5238 gcccttctgtcttctttagaggtgaggaacaccccttgacagcagcagcagcagcagcagcagc 5297
 QY 2361 attgctcttaaggggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2420
 Db 5298 attgctcttaaggggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5357
 QY 2421 gctgcttcccaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2480
 Db 5358 gctgcttcccaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5417
 QY 2481 gtaggaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2540
 Db 5418 gtaggaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5477

QY 2541 agltgagatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2600
 Db 5478 agltgagatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5537
 QY 2601 gggacactatctcgtactctttagac 2660
 Db 5538 gggacactatctcgtactctttagac 5597
 QY 2661 ctgattccatgcatcccaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2686
 Db 5598 ctgattccatgcatcccaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5623

RESULT 4

AA77726 standard, DNA; 7333 BP.

AA77726;

23-MAY-2001 (first entry)

Defective retroviral genome isolated from PK-15 cell line.

Retrovirus; graft transplantation; xenotransplantation; PK-15 cell line;

ss.

Unidentified.

US6190861-B1.

20-FEB-2001.

13-DEC-1996; 96US-0766528.

14-DEC-1995; 95US-0572645.

(GEHO) GEN HOSPITAL CORP.

Fishman JA.

WPI, 2001-256211/26.

P-PSDB; AAB73282, AAB73283, AAB73284.

Assessing risk of endogenous retroviruses in clinical practice and in

xenotransplantation, comprises using probe sequences derived from swine

or miniature swine retroviral genome.

Claim 1; Fig 2; 127pp; English.

The present invention relates to a method for screening a cell or tissue

for the presence or expression of a retrovirus (RV), comprising

contracting a target nucleic acid from the cell or tissue with a second

nucleic acid (first). The present invention (e.g. the present sequence or a

fragment thereof). The method is useful for RV detection and to assess

graft transplantation risk. Screening of animals allows the elimination

of donors with active replication of known viruses. Inactive retroviruses

can be detected and inactivated, allowing identification and elimination

of potential human pathogens derived from swine in a manner not possible

in the outbred human organ donor population and is important to the

development of human xenotransplantation.

Sequence 7333 BP; 1984 A; 1788 C; 1885 G; 1676 T; 0 other;

Query Match 32.9%; Score 1091; DB 22; Length 7333;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 1291; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1364 aaaggggttgccttaccctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1423

Db 3517 aaaggggttgccttaccctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3576

QY 1424 tatagaagccttaccattgccaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1483

```

|||||
Db 3577 tattagaagccttacatttgccaaaaggctagctatataacatgctctctgacatcaga 3636
QY 1484 aagccaaagatctcatatctagagggaaccagatgctgacccgggtgtccaaagcgagcag 1543
Db 3637 aagccaaagatctcatatctagagggaaccagatgctgacccgggtgtccaaagcgagcag 3696
QY 1544 ccagagctgttaacctctctgacatataagaaacgcccacaaagcccccagacacagac 1603
Db 3697 ccagagctgttaacctctctgacatataagaaacgcccacaaagcccccagacacagac 3756
QY 1604 agtacaccctagaagactgagcaagagataaaaaagatagacagatctctgagactcgg 1663
Db 3757 agtacaccctagaagactgagcaagagataaaaaagatagacagatctctgagactcgg 3816
QY 1664 aggggacctgtctatctacatataaggaagaaatctctgcccacaaagaaaggttagaat 1723
Db 3817 aggggacctgtctatctacatataaggaagaaatctctgcccacaaagaaaggttagaat 3876
QY 1724 atgtccaaacagatatacgtlctaacaccctagaaactaaacacctgcagcagttgttca 1783
Db 3877 atgtccaaacagatatacgtlctaacaccctagaaactaaacacctgcagcagttgttca 3936
QY 1784 gaacatcccttataatgttctgagagctaccagaggtgctgactcgtgtgttcaaacatt 1843
Db 3937 gaacatcccttataatgttctgagagctaccagaggtgctgactcgtgtgttcaaacatt 3996
QY 1844 gtgtgcccctgcagcctggttaatgataccttcacgaatacctccaggaagagactaa 1903
Db 3997 gtgtgcccctgcagcctggttaatgataccttcacgaatacctccaggaagagactaa 4056
QY 1904 ggggaaagccaaacccagcgcctactgaggaagtgtgacttcaatgaagtaaacgcggtaaat 1963
Db 4057 ggggaaagccaaacccagcgcctactgaggaagtgtgacttcaatgaagtaaacgcggtaaat 4116
QY 1964 acggaacaaatactatgttgtttgttagacaccttccagagatgtgtagaggtctaac 2023
Db 4117 acggaacaaatactatgttgtttgttagacaccttccagagatgtgtagaggtctaac 4176
QY 2024 ctactaagaagaagacttcaacccgtgtgctlaagaaatactgtaggaagaattttccaa 2083
Db 4177 ctactaagaagaagacttcaacccgtgtgctlaagaaatactgtaggaagaattttccaa 4236
QY 2084 gatttggaaatacctaaggttaataaggttcagacaatgtgtccagcttctgtgccaggtaa 2143
Db 4237 gatttggaaatacctaaggttaataaggttcagacaatgtgtccagcttctgtgccaggtaa 4296
QY 2144 gtaaggaactggtccaaagatactgggagatgtggaactgtgtgcatatacagacccc 2203
Db 4297 gtaaggaactggtccaaagatactgggagatgtggaactgtgtgcatatacagacccc 4356
QY 2204 aaagctcagaagcagtagagaagagatgaatagaacccaataaagaagacccttaaccaatga 2263
Db 4357 aaagctcagaagcagtagagaagagatgaatagaacccaataaagaagacccttaaccaatga 4416
QY 2264 ccagagagactgtgcatataagatgtgattgtgctctctgaccttgcctttagaggtga 2323
Db 4417 ccagagagactgtgcatataagatgtgattgtgctctctgaccttgcctttagaggtga 4476
QY 2324 ggaacacccctgagcagtttgggtctgacccctatgaatgtctcctcaggggagaccccc 2383
Db 4477 ggaacacccctgagcagtttgggtctgacccctatgaatgtctcctcaggggagaccccc 4536
QY 2384 cgttggcaagaatgtgcttgcacatagtgatgtgctgtcttccagccttgttct 2443
Db 4537 cgttggcaagaatgtgcttgcacatagtgatgtgctgtcttccagccttgttct 4596
QY 2444 ctgagctcaagcgcctcagatgtgtgagcagcagcgtgtggaagcagctccggagagcct 2503
Db 4597 ctgagctcaagcgcctcagatgtgtgagcagcagcgtgtggaagcagctccggagagcct 4656
QY 2504 acccagagagagactgtgcaagttccacatgccttccaatttgaagattagttatgtta 2563

```

```

Db 4657 acccagagagagactgtgcaagttccacatgccttccaaagttgagatcagttatgtta 4716
QY 2564 gacgccaccgtgcaggaagaaacctcgagactcgtgtgaagggagacctaactcgtactttga 2623
Db 4717 gacgccaccgtgcaggaagaaacctcgagactcgtgtgaagggagacctaactcgtactttga 4776
QY 2624 ccacaccaacgctgtgtgaagatcgaaggaatcccc 2658
Db 4777 ccacaccaacgctgtgtgaagatcgaaggaatcccc 4811

RESULT 5
AAT74883
ID AAT74883 standard; cDNA: 7393 BP.
XX
AC AAT74883;
XX
DT 09-FEB-1998 (first entry)
XX
DE Porcine retrovirus cDNA (defective).
XX
KW Retrovirus; porcine; GAG protein; POL protein; ENV protein;
KW xenotransplantation; infectious; provirus; organ transplant; donor;
KW activated virus; PCR; ss.
XX
OS Porcine retrovirus.
XX
FH Key
FH CDS Location/Qualifiers
FT 598..2172
FT /*tag= a
FT 598..2169
FT /*tag= b
FT /*note= "putative GAG protein"
FT CDS 2320..4737
FT /*tag= c
FT /*note= "putative POL coding region (partial) as
FT described in the specification"
FT
FT mat_peptide 2320..3522
FT /*tag= d
FT /*note= "putative POL protein (partial)"
FT 3516..4328
FT /*tag= e
FT /*note= "putative POL protein (partial)"
FT CDS 4332..4748
FT /*tag= f
FT 4738..6725
FT /*tag= g
FT /*note= "putative ENV coding region (partial) as
FT described in the specification"
FT mat_peptide 4752..6722
FT /*tag= h
FT /*note= "ENV protein (partial)"

W09721836-A1.
19-JUN-1997.
XX
PF 13-DEC-1996; 96WO-US19680.
XX
PR 14-DEC-1995; 95US-0572645.
XX
PA (GEHO ) GEN HOSPITAL CORP.
XX
PI Fishman JA;
XX
DR WPI: 1997-332804/30.
DR P-PSDB: AAM32091-W32095.
XX
PT New nucleic acid from porcine retroviruses - used for detecting
PT viruses in transplant or other tissue and for assessing risk of
PT transmitting infection to graft recipient
XX
PS Claim 16; Fig 2; 128bp; English.

```

XX This cDNA sequence represents a defective purified swine retrovirus
CC found in PK-15 cells containing the putative coding regions for viral
CC GAg, POL and ENV proteins. There are a few in frame stop codons and
CC apparent frame shifts in the given coding sequence which alter features
CC of the translation. This sequence and PCR fragments generated from the
CC sequence (see AAT74812-74882) could be used to screen organs for the
CC presence of porcine retroviruses prior to xenotransplantation.
CC Transplantation can increase the likelihood of retroviral activation if
CC intact and infectious proviruses are present. The porcine retroviral
CC sequence can be used to generate probes to determine the level (e.g.
CC copy number) of intact (i.e. potentially replicating) porcine provirus
CC sequences in a strain of xenograft transplantation donors. It can be
CC used to detect mutations, genetic lesions or viral recombinants and to
CC determine the histological localisation of activated retroviruses. Using
CC Polymerase Chain Reaction DNA Quantitation (PQD) on blood mononuclear
CC cells, infectivity titration and susceptibility testing can be
CC performed. Ultimately animal donors without intact porcine retroviral
CC sequences or a lower copy number of viral elements could be selected.
XX

Sequence 7393 BP; 2004 A; 1801 C; 1896 G; 1692 T; 0 other;

Query Match	32.9%	Score 1091;	DB 18;	Length 7393;
Best Local Similarity	99.7%;	Pred. No. 0;		
Matches 1291; Conservative	0;	Mismatches 4;	Indels 0;	Gaps 0;

QY	1364	aagggggtgtgtcttacctagcagggagggagaataaaagaacaagaggaattcttaagcc	1423
Db	3517	aaagggggtgtgtcttacctagcagggagggagaataaaagaacaagaggaattcttaagcc	3576
QY	1424	tattagaagccttacattgtgccaaaagggcagctattatcacctgtgcccgtgacatcaga	1483
Db	3577	tattagaagccttacattgtgccaaaagggcagctattatcacctgtgcccgtgacatcaga	3636
QY	1484	aagccaagaatcctcatatcattagaggagaaaccagatgtgtgacccgggtttgccaaacagcgag	1543
Db	3637	aagccaagaatcctcatatcattagaggagaaaccagatgtgtgacccgggtttgccaaacagcgag	3696
QY	1544	ccgaagctgtttaaccttctgtcctataatagaagaagccccaagcccgaagcccgaagccgaagcagc	1603
Db	3697	cccaagctgtttaaccttctgtcctataatagaagaagccccaagcccgaagcccgaagccgaagcagc	3756
QY	1604	agcacaccccttagaagaactgtgcagaagatataaaaaagatagacagattctcttgagaacccgag	1663
Db	3757	agcacaccccttagaagaactgtgcagaagatataaaaaagatagacagattctcttgagaacccgag	3816
QY	1664	aagggaacctgtctatacctcatatggygaagaaatctctgcccacaagaagaggtttagaat	1723
Db	3817	aagggaacctgtctatacctcatatggygaagaaatctctgcccacaagaagaggtttagaat	3876
QY	1724	atgtccaagaagttatattgttcttaacccaacccataggaactataaacctgtgcagcagttgtgtca	1783
Db	3877	atgtccaagaagttatattgttcttaacccaacccataggaactataaacctgtgcagcagttgtgtca	3936
QY	1784	gaacatcccccattacatgtttcttgaggtctaccagaagtggtgtgactcogtvggtcaaacatt	1843
Db	3937	gaacatcccccattacatgtttcttgaggtctaccagaagtggtgtgactcogtvggtcaaacatt	3996
QY	1844	gtgtgcctgtgcagctgtgttaatgtcctaactcctccagaatlaactccaggaagaagagactaa	1903
Db	3997	gtgtgcctgtgcagctgtgttaatgtcctaactcctccagaatlaactccaggaagaagagactaa	4056
QY	1904	gggggaagccaaccagcgctacatcgggaagaagtggaacttcattggtgtaaaagccggctcaaat	1963
Db	4057	gggggaagccaaccagcgctacatcgggaagaagtggaacttcattggtgtaaaagccggctcaaat	4116
QY	1964	acggaacaacatatcatattgttttttgtagaaccttttcagaatgtggtagaggtctatc	2023
Db	4117	acggaacaacatatcatattgttttttgtagaaccttttcagaatgtggtagaggtctatc	4176
QY	2024	ctactcagaagaagagacttcaacgctgtgtgtcgttaagaanaatactgtgaggaatttttccaa	2083

Db	4177	caactaaaaaagagacttcaaccgttggtgctagaagaataactcggaggaatttttcca	42305			
Qy	2084	gatttggatacctaagaataagagtcacagacaatggtccagcttcgttcccgatga	2143			
Db	4237	gatttggataccctaagtaataaggtgcagacaatgtccagcttcgttcccgatga	4296			
Qy	2144	gtcagggaactcggccaagaatatcttgggatttgatttggaaactcgtcaattgtgcatacagaacc	2203			
Db	4297	gtcagggaactcggccaagaatatcttgggatttgatttggaaactcgtcaattgtgcatacagaacc	4356			
Qy	2204	aaagctcagggacggttagagaagatgaataagaaccattaaagaagacccttaacaaattga	2265			
Db	4357	aaagctcagggacggttagagaagatgaataagaaccattaaagaagacccttaacaaattga	4416			
Qy	2264	ccacaagagactcgtgattaaagatcttgatggtctctccctcgtccttctgctttaaagtga	2323			
Db	4417	ccacaagagactcgtgattaaagatcttgatggtctctccctcgtccttctgctttaaagtga	4476			
Qy	2324	ggaacaccccttgacagtttgggctgaacccctatgaattgtccttaaggggagacccccc	2383			
Db	4477	ggaacaccccttgacagtttgggctgaacccctatgaattgtccttaaggggagacccccc	4536			
Qy	2384	cgtttggcagaagaattgcctttgacacatagatgctgatgtgtctgtcttccagccttgttct	2443			
Db	4537	cgtttggcagaagaattgcctttgacacatagatgctgatgtgtctgtcttccagccttgttct	4596			
Qy	2444	ctagagcccaaggcgcgtcgagtggttgaggcagcagcagcgtggagagcagctcccgaggagcct	2503			
Db	4547	ctagagcccaaggcgcgtcgagtggttgaggcagcagcagcgtggagagcagctcccgaggagcct	4566			
Qy	2504	actcagaagagagactctgcaaatctccacatcgccttccaaagttagaagatcagtcatagtga	2563			
Db	4657	actcagaagagagactctgcaaatctccacatcgccttccaaagttagaagatcagtcatagtga	4716			
Qy	2564	gagcccaacgcgtgcaagaaacctcgagaccccggttggaaggagacctatctgtactttga	2623			
Db	4717	gagcccaacgcgtgcaagaaacctcgagaccccggttggaaggagacctatctgtactttga	4776			
Qy	2624	ccacaccaaagcgtctggaagtgcgaaggaatccccc	2658			
Db	4777	ccacaccaaagcgtctggaagtgcgaaggaatccccc	4811			
RESULT 6						
AAC67020						
ID	AAC67020 standard; DNA; 4402 BP.					
AC	AAC67020;					
XX						
DT	27-MAR-2001 (first entry)					
XX						
DE	PERV env protein coding sequence SEQ ID NO: 20.					
XX						
KW	Xenotransplantation; infectious agent; vaccine; ds.					
XX						
OS	Porcine endogenous retrovirus.					
XX						
PN	WO200071726-A1.					
XX						
PD	30-NOV-2000.					
XX						
PF	24-MAY-2000; 2000WO-US14296.					
XX						
PR	24-MAY-1999; 99US-0135631.					
XX						
PA	(MAYO-) MAYO MEDICAL VENTURES.					
PI	Federspiel MJ;					
XX						
DR	WPI; 2001-032041/04.					
XX						
Inhibiting or preventing infectious agent transmission in mammalian						
transplant recipients, by introducing recombinant DNA comprising DNA						

PT encoding extracellular proteins of the agent into donor cells, such as
 PT swine cells -
 XX
 PS Claim 16, Page 105-106; 144pp; English.
 XX
 CC The present invention provides a method to prevent the transmission of
 CC infectious agents during xenotransplantation. This involves introducing
 CC to donor swine cells a recombinant DNA encoding a peptide fragment from
 CC the infectious agent, and then introducing these cells into the
 CC transplant recipient.
 XX
 SQ Sequence 4402 BP; 1259 A; 1085 C; 1111 G; 947 T; 0 other;

Query Match 25.5%; Score 848; DB 22; Length 4402;
 Best Local Similarity 98.8%; Pred. No. 0;
 Matches 2198; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 31 ttctgctgagatgaacccactagcaacacatttgccttcgaatgagagatcca 90
 Db 1043 ttctgctgagatgaacccactagcaacacatttgccttcgaatgagagatcca 1102
 QY 91 ggtacgggaagaacccgggagctcactgagccgagctgcccagaaggttcaagaatcc 150
 Db 1103 ggtacgggaagaacccgggagctcactgagccgagctgcccagaaggttcaagaatcc 1162
 QY 151 ccgacacatcttgacgaagccctacacagagaccttgcccaacttcagatccacaccc 210
 Db 1163 ccgacacatcttgacgaagccctacacagagaccttgcccaacttcagatccacaccc 1222
 QY 211 caggtgacccctctccagatcgttggatgacctctcttcgcygagccacacacagac 270
 Db 1223 caggtgacccctctccagatcgttggatgacctctcttcgcygagccacacacagac 1282
 QY 271 tgcctaaaggtacgaagagcactacgtctgaaatgtctactacgaactcaagagcctc 330
 Db 1283 tgcctaaaggtacgaagagcactacgtctgaaatgtctactacgaactcaagagcctc 1342
 QY 331 gctaaagaagccacagatttcagagagagagtaacatacttgggtacagtttcgcygac 390
 Db 1343 gctaaagaagccacagatttcagagagagagtaacatacttgggtacagtttcgcygac 1402
 QY 391 gggcagcagatgctcagcagagcagcagaaagaacttgaatccagataccggcccaacc 450
 Db 1403 gggcagcagatgctcagcagagcagcagaaagaacttgaatccagataccggcccaacc 1462
 QY 451 acacccaacaagaatgagagagtttttgggagcagctgagatttcgacagatgtgacccg 510
 Db 1463 acacccaacaagaatgagagagtttttgggagcagctgagatttcgacagatgtgacccg 1522
 QY 511 gggtttgcagacttagcagcccaactacccgtaacccaagaaggggggattctcc 570
 Db 1523 gggtttgcagacttagcagcccaactacccgtaacccaagaaggggggattctcc 1582
 QY 571 tgggctctcagcagcagagagcatttgcgtatataaagaagccctgcgagcagaccc 630
 Db 1583 tgggctctcagcagcagagagcatttgcgtatataaagaagccctgcgagcagaccc 1642
 QY 631 gctctggccctccctgacgtaactaaaccccttaacccctatctatgtgagtgagtaagga 690
 Db 1643 gctctggccctccctgacgtaactaaaccccttaacccctatctatgtgagtgagtaagga 1702
 QY 691 gtagccggagagattttaacccaacccctagagacatgtagagagagacccgttgcctactg 750
 Db 1703 gtagccggagagattttaacccaacccctagagacatgtagagagagacccgttgcctactg 1762
 QY 751 tcaagaagagcttgatctctgtagccagtggttggccgtatgctggaaggtatccgagct 810
 Db 1763 tcaagaagagcttgatctctgtagccagtggttggccgtatgctggaaggtatccgagct 1822
 QY 811 gtggcctactggtcgaagagcgtgacaatgtgacttggagacgaataataactgtatata 870
 Db 1823 gtggcctactggtcgaagagcgtgacaatgtgacttggagacgaataataactgtatata 1882

QY 871 gcccccacatgtagagagacatcgttcgacagcccccagaccgattgatatcaacgccc 930
 Db 1883 gcccccacatgtagagagacatcgttcgacagcccccagaccgattgatatcaacgccc 1942
 QY 931 cgcagtagccacatcaacagcctgtcttcaacagagaggtgacatttcgctccacagccc 990
 Db 1943 cgcagtagccacatcaacagcctgtcttcaacagagaggtgacatttcgctccacagccc 2002
 QY 991 gcttcaacccctgcacactctctgcttgaagaagatgtatgaacagctgactatgatgac 1050
 Db 2003 gcttcaacccctgcacactctctgcttgaagaagatgtatgaacagctgactatgatgac 2062
 QY 1051 calcaactatgtatgagagagcttgggttcgcaagagaccccttaacagacatccgtgact 1110
 Db 2063 calcaactatgtatgagagagcttgggttcgcaagagaccccttaacagacatccgtgact 2122
 QY 1111 ggaagaaggtcacttgccttcaacttgcaggaagcagctatgtgtgtgaaagttaagaagatg 1170
 Db 2123 ggaagaaggtcacttgccttcaacttgcaggaagcagctatgtgtgtgaaagttaagaagatg 2182
 QY 1171 gcttgggcygagtggttggacggygaccccgacagatcttggccagagccttgcggaagga 1230
 Db 2183 gcttgggcygagtggttggacggygaccccgacagatcttggccagagccttgcggaagga 2242
 QY 1231 acttaagcgcaaaaggtcgtagctatgtgcctcaacgaagcttgcggtcgcggaaggg 1290
 Db 2243 acttaagcgcaaaaggtcgtagctatgtgcctcaacgaagcttgcggtcgcggaaggg 2302
 QY 1291 aatccataaacttatacgcagacagagatgtccttgcagatgcacagatcaacggy 1350
 Db 2303 aatccataaacttatacgcagacagagatgtccttgcagatgcacagatcaacggy 2362
 QY 1351 gccatctataacaaaggggttgccttaacctcaacgagggaggaataaagaacaagag 1410
 Db 2363 gccatctataacaaaggggttgccttaacctcaacgagggaggaataaagaacaagag 2422
 QY 1411 gaaattctaaagcctattgaaagccttatacttggccaaaaggtgacattatatacgt 1470
 Db 2423 gaaattctaaagcctattgaaagccttatacttggccaaaaggtgacattatatacgt 2482
 QY 1471 ccttgacatcagaagaagccaaagatctatctatgaagaagaaacagatgctgcagcgggt 1530
 Db 2483 ccttgacatcagaagaagccaaagatctatctatgaagaagaaacagatgctgcagcgggt 2542
 QY 1531 gccaaagcagcagcccaaggttgccttaaccttgcctataatagaacacgcccacaccca 1590
 Db 2543 gccaaagcagcagcccaaggttgccttaaccttgcctataatagaacacgcccacaccca 2602
 QY 1591 gaacccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1650
 Db 2603 gaacccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2662
 QY 1651 tctgaagcctcgagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1710
 Db 2663 tctgaagcctcgagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2722
 QY 1711 gaaaggttagaatgttccaacagatatactgcttaacccactagaagaaactaacaactg 1770
 Db 2723 gaaaggttagaatgttccaacagatatactgcttaacccactagaagaaactaacaactg 2782
 QY 1771 cagcaggttgcagaaacatccctctatcaatgttctggaagcctacagagagtggtgcagc 1830
 Db 2783 cagcaggttgcagaaacatccctctatcaatgttctggaagcctacagagagtggtgcagc 2842
 QY 1831 gtgtgtaaacatctgtgtgcctgcagcagctgtgttaattgtaactcttccagaataactcca 1890
 Db 2843 gtgtgtaaacatctgtgtgcctgcagcagctgtgttaattgtaactcttccagaataactcca 2902
 QY 1891 ggaagaagactaaggggaagccacacagcgtctacgtgggaagtgtgacttcaactgagta 1950
 Db 2903 ggaagaagactaaggggaagccacacagcgtctacgtgggaagtgtgacttcaactgagta 2962

QY 1951 aagccggtactaatacaggaacaatatctatgtgtttttagagacacttttcagatcg 2010
|||||
Db 2963 aagccggtactaatacaggaacaatatctatgtgtttttagagacacttttcagatcg 3022
QY 2011 gtaggggtctatctactaagaagaagacttcaacctggtggtgctaaagaatactatcg 2070
|||||
Db 3023 gtaggggtctatctactaagaagaagacttcaacctggtggtgctaaagaatactatcg 3082
QY 2071 gaaattttccagaatttgaatacctaaggaaggaaggtcagacaaggttccacattc 2130
|||||
Db 3083 gaaattttccagaatttgaatacctaaggaaggaaggtcagacaaggttccacattc 3142
QY 2131 gttgccaggttaagtcagggactggtccaaagatatgtgggattgttggaactgcatgt 2190
|||||
Db 3143 gttgccaggttaagtcagggactggtccaaagatatgtgggattgttggaactgcatgt 3202
QY 2191 gatacagagcccaagagtcagagacagtgtagagagatgaataagacaatcaagaagcc 2250
|||||
Db 3203 gatacagagcccaagagtcagagacagtgtagagagatgaataagacaatcaagaagcc 3262
QY 2251 ctac 2255
|||||
Db 3263 ctac 3267
RESULT 7
AAC67021
ID AAC67021 standard; DNA; 6076 BP.
XX AAC67021;
AC AAC67021;
DT 27-MAR-2001 (first entry)
XX PERV env protein coding sequence SEQ ID NO: 21.
DE
XX Xenotransplantation; infectious agent; vaccine; ds.
KW Porcine endogenous retrovirus.
OS
XX WO200071726-A1.
PN
XX 30-NOV-2000.
PD
XX 24-MAY-2000; 2000WO-US14296.
PF
XX 24-MAY-1999; 99US-0135631.
PR
XX (MAYO-) MAYO MEDICAL VENTURES.
PA
XX Federspiel MJ;
PI
XX WPI; 2001-032041/04.
DR
XX
XX Inhibiting or preventing infectious agent transmission in mammalian
PT Transplant recipients, by introducing recombinant DNA comprising DNA
PT encoding extracellular proteins of the agent into donor cells, such as
PT swine cells -
PS
XX Claim 16; Page 107-109; 144pp; English.
XX
XX The present invention provides a method to prevent the transmission of
CC infectious agents during xenotransplantation. This involves introducing
CC to donor swine cells a recombinant DNA encoding a peptide fragment from
CC the infectious agent, and then introducing these cells into the
CC transplant recipient.
XX
XX Sequence 6076 BP; 1613 A; 1512 C; 1609 G; 1342 T; 0 other;

Query Match 24.0%; Score 797; DB 22; Length 6076;
Best Local Similarity 98.7%; Pred. No. 3.2e-314;
Matches 2197; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 31 tctgcctgaagattacacccaccactagccaacacattttgcctcgaatgagatcca 90
|||||
Db 3577 tctgcctgaagattacacccaccactagccaacacattttgcctcgaatgagatcca 3636
QY 91 ggtacgggaagaacccgggacgttcaacttgacccgactgcccgaagggtccaagaactcc 150
|||||
Db 3637 ggtacgggaagaacccgggacgttcaacttgacccgactgcccgaagggtccaagaactcc 3696
QY 151 ccgagccctctttgcgaagagcccttacacagggagacttgccaactttcagatccaacacct 210
|||||
Db 3697 ccgagccctctttgcgaagagcccttacacagggagacttgccaactttcagatccaacacct 3756
QY 211 caagttgacccctccctccagtaacgttgatgacctgtcttcgcygagagccaccaacaagac 270
|||||
Db 3757 caagttgacccctccctccagtaacgttgatgacctgtcttcgcygagagccaccaacaagac 3816
QY 271 tgccttaagaagttacgaagagcaactacgtcgtggaattgtctgaactaagctacagacct 330
|||||
Db 3817 tgccttaagaagttacgaagagcaactacgtcgtggaattgtctgaactaagctacagacct 3876
QY 331 gctaaagaagcccccagatttcagagaagaaggttaacatacttgggttacaatttgcgggac 390
|||||
Db 3877 gctaaagaagcccccagatttcagagaagaaggttaacatacttgggttacaatttgcgggac 3936
QY 391 gggcagcgatggtcgaagcagagcagcaggaagaagaactgtaaccagataccgcccacac 450
|||||
Db 3937 gggcagcgatggtcgaagcagagcagcaggaagaagaactgtaaccagataccgcccacac 3996
QY 451 acagccaacaagaatgagagaagttttggggacagcttggaatttcgaagacttggatcccg 510
|||||
Db 3997 acagccaacaagaatgagagaagttttggggacagcttggaatttcgaagacttggatcccg 4056
QY 511 ggggttgcagccttagacagcccccacttaccgcttaacccaagaagaaggaggaattctcc 570
|||||
Db 4057 ggggttgcagccttagacagcccccacttaccgcttaacccaagaagaaggaggaattctcc 4116
QY 571 tgggtctctgaagcaccgaagaagcatltgactgctatacaaaaaggccctgctgaagcagcct 630
|||||
Db 4117 tgggtctctgaagcaccgaagaagcatltgactgctatacaaaaaggccctgctgaagcagcct 4176
QY 631 gctctgacctccctgaagcgttaactaaaccttacccttaagtgtgagtgaagcgtlaagg 690
|||||
Db 4177 gctctgacctccctgaagcgttaactaaaccttacccttattgtgagtgaagcgtlaagg 4236
QY 691 gtagcccggaagatgtttaaccacaaccttagagacatgtagagagactgttgcctacccg 750
|||||
Db 4237 gtagcccggaagatgtttaaccacaaccttagagacatgtagagagactgttgcctacccg 4296
QY 751 tcaagaagacttgatccctgtagccagtggttggccgtaagtctgaaggtatcgagct 810
|||||
Db 4297 tcaagaagacttgatccctgtagccagtggttggccgtaagtctgaaggtatcgagct 4356
QY 811 gtagccactggtcgaagagcgtgtaacatltgactttggagacagaataactgtaata 870
|||||
Db 4357 gtagccactggtcgaagagcgtgtaacatltgactttggagacagaataactgtaata 4416
QY 871 gcccoccatgcatltgagaacatcgtttggccaccccccagaccgagatgtagaccagcc 930
|||||
Db 4417 gcccoccatgcatltgagaacatcgtttggccaccccccagaccgagatgtagaccagcc 4476
QY 931 cgaatgacccatacaaaagcctgctcttcacagagaagggtaccttgcctccaccagcc 990
|||||
Db 4477 cgaatgacccatacaaaagcctgctcttcacagagaagggtacgttgcctccaccagcc 4536
QY 991 gctctcaaccctgacactcttgcctgaagagactgtagaacagtgatgaatgctatgctc 1050
|||||
Db 4537 gctctcaaccctgacactcttgcctgaagagactgtagaacagtgatgaatgctatgctc 4596
QY 1051 catcaactatgtattgagagagactgggtccggaagaagacttaacagacatacgcgtgact 1110
|||||
Db 4597 catcaactatgtattgagagagactgggtccggaagaagacttaacagacatacgcgtgact 4656
QY 1111 gtagaagtgtaactcgtgttcaactgaagagcagatgtgtgtgagaagtaagaagatg 1170


```
D 3882 tgccttagaaggcaggaagcactactctgtaattgtctgcactaggtacacagagctct 3941
Q 331 gctaaagaagccagagattgcaagagaaggttaacactcttggtgtacagtttgaggc 390
D 3942 gctaaagaagccagagattgcaagagaaggttaacactcttggtgtacagtttgaggc 4001
Q 391 gggcagcgtatgctgagcaggaagcaggaagaaactgtatccagatataccgcccacac 450
D 4002 gggcagcgtatgctgagcaggaagcaggaagaaactgtatccagatataccgcccacac 4061
Q 451 aacagccaaacaaagttagaagaatttttgggagacagcttgatcttgcagactgtgaccg 510
D 4062 aacagccaaacaaagttagaagaatttttgggagacagcttgatcttgcagactgtgaccg 4121
Q 511 gggattgagaccttaagcagcccaactctacccgttaacaaagaagaaagggagattccc 570
D 4122 gggattgagaccttaagcagcccaactctacccgttaacaaagaagaaagggagattccc 4181
Q 571 tgggtctctgagcagcaggaagcatttgaatgctatcaaaaaggccctgtgagcgacac 630
D 4182 tgggtctctgagcagcaggaagcatttgaatgctatcaaaaaggccctgtgagcgacac 4241
Q 631 gctctggacctccctggagcttaactaaaccttacccttattgtgtatgagcgtaaggga 690
D 4242 gctctggacctccctggagcttaactaaaccttacccttattgtgtatgagcgtaaggga 4301
Q 691 gtagcccgagagatttaacccaacccctagagaccatgagagaagccgtgtgacctg 750
D 4302 gtagcccgagagatttaacccaacccctagagaccatgagagaagccgtgtgacctg 4361
Q 751 tcaaaagaagcttgatccctgttagcagatgtgtgcccgtatgtctgaagctatccgacgt 810
D 4362 tcaaaagaagcttgatccctgttagcagatgtgtgcccgtatgtctgaagctatccgacgt 4421
Q 811 gtagccatactgtgtaagaagcgttgaacaaattgacttgggacagaaataataactgtata 870
D 4422 gtagccatactgtgtaagaagcgttgaacaaattgacttgggacagaaataataactgtata 4481
Q 871 gcccccactgcatctggaagaacatctgtcgcagccccagacagctgaatgacaaagccc 930
D 4482 gcccccactgcatctggaagaacatctgtcgcagccccagacagctgaatgacaaagccc 4541
Q 931 cgcatagaccactatcaaaagcctgtctctcaagaagaaggtcactctgcctcaacagcc 990
D 4542 cgcatagaccactatcaaaagcctgtctctcaagaagaaggtcactctgcctcaacagcc 4601
Q 991 gctcacaacccctgcaacccctctctgcctgaagaagcgtgtagaagcaggtacccatgattgc 1050
D 4602 gctcacaacccctgcaacccctctctgcctgaagaagcgtgtagaagcaggtacccatgattgc 4661
Q 1051 catcaactattgattgagagagactggtgtccgcaagagccttacagacataccgctgact 1110
D 4662 catcaactattgattgagagagactggtgtccgcaagagccttacagacataccgctgact 4721
Q 1111 ggaagaatgctaaacctgttctactgacgggaagcagctatgtgtggaaggttaagagatg 1170
D 4722 ggaagaatgctaaacctgttctactgacgggaagcagctatgtgtggaaggttaagagatg 4781
Q 1171 gctggggcgagctgtgtgagcggagcccgacgactgtggccagacagccctgcgggaagga 1230
D 4782 gctggggcgagctgtgtgagcggagcccgacgactgtggccagacagccctgcgggaagga 4841
Q 1231 acttcagcgcaaaaagcgtgagctatgacctcagcgaacacttgcggctgcggcaaggg 1290
D 4842 acttcagcgcaaaaagcgtgagctatgacctcagcgaacacttgcggctgcggcaaggg 4901
Q 1291 aaatccataaacttatacggagacaggtatgaccttgcgactgcacagctacacggg 1350
D 4902 aaatccataaacttatacggagacaggtatgaccttgcgactgcacagctacacggg 4961
Q 1351 gccatctataaacaaggggttgcttaactcagcaggggggaaataaagaagaacaag 1410
D 4962 gccatctataaacaaggggttgcttaactcagcaggggggaaataaagaagaacaag 5021

Q 1411 gaattctaagccttattgaagccttaccattgcccataaagaagctagctattatacactgt 1470
D 5022 gaattctaagccttattgaagccttaccattgcccataaagaagctagctattatacactgt 5081
Q 1471 cctggacatcagaaagccaaagatctatctagagaggaacccagatggtgcagccgggtc 1530
D 5082 cctggacatcagaaagccaaagatctatctagagaggaacccagatggtgcagccgggtc 5141
Q 1531 gccaaagcagcagcccaagcgtgttaacccctctgctatataatagaagaacgcccacccca 1590
D 5142 gccaaagcagcagcccaagcgtgtgttaacccctctgctatataatagaagaacccca 5201
Q 1591 gaaccacaagcagctacacccctgaagaagctgacagaagataaaaagatagaccagttc 1650
D 5202 gaaccacaagcagctacacccctgaagaagctgacagaagataaaaagatagaccagttc 5261
Q 1651 tctgagactccgggagggagctgtgctatatacctcatatggaaggaaatccctgcccacaaa 1710
D 5262 tctgagactccgggagggagctgtgctatatacctcatatggaaggaaatccctgcccacaaa 5321
Q 1711 gaaaggttagaataatgtccaaacagatacatcgtctaaacccacccctaggaactaaacactg 1770
D 5322 gaaaggttagaataatgtccaaacagatacatcgtctaaacccacccctaggaactaaacactg 5381
Q 1771 cagcaagttgttcagaaacatccctctatcatatgcttctgaagctaccaggaagtgctgactg 1830
D 5382 cagcaagttgttcagaaacatccctctatcatatgcttctgaagctaccaggaagtgctgactg 5441
Q 1831 gtgttcaacaactgtgtgcccctgcccagctgtgtaagttaatccctccaaataaccctca 1890
D 5442 gtgttcaacaactgtgtgcccctgcccagctgtgtaagttaatccctccaaataaccctca 5501
Q 1891 ggaagaagcactaaagggaagccacccagcgtcactggtggaagtggaacttgcactgagta 1950
D 5502 ggaagaagcactaaagggaagccacccagcgtcactggtggaagtggaacttgcactgagta 5561
Q 1951 aagccggcctaataagcgaacaaatatctatgtgttttgaacacaccttcaagatg 2010
D 5562 aagccggcctaataagcgaacaaatatctatgtgttttgaacacaccttcaagatg 5621
Q 2011 gtagagcttatactactaagaagaagagacttcaacggtgtgtgctaaagaataactggag 2070
D 5622 gtagagcttatactactaagaagaagagacttcaacggtgtgtgctaaagaataactggag 5681
Q 2071 gaaattttccaagatttggaaataactaagaaggtataggttcagacaatggtccagcttc 2130
D 5682 gaaattttccaagatttggaaataactaagaaggtataggttcagacaatggtccagcttc 5741
Q 2131 gttggccaggttaagctcagggagactggtccaaagatatgttggtatgtggaactgactgt 2190
D 5742 gttggccaggttaagctcagggagactggtccaaagatatgttggtatgtggaactgactgt 5801
Q 2191 gcatlacagaccccaagcctcagagcaggttagaagaaggtatgaatgaacccattaaagacc 2250
D 5802 gcatlacagaccccaagcctcagagcaggttagaagaaggtatgaatgaacccattaaagacc 5861
Q 2251 cttac 2255
D 5862 cttac 5866

RESULT 9
AAV82748
ID AAV82748 standard; DNA; 2462 BP.
XX
XX AAV82748;
AC
XX
XX 25-FEB-1999 (first entry)
DT
XX
XX Pig endogenous retrovirus (PERV)-A envelope (env) gene region.
DE
XX Pig endogenous retrovirus; PERV-A; envelope protein; PERV-B; subtype;
KW
```


KW probe: primer; detection: retrovirus; human tissue; xenotransplant;
 KW primary porcine tissue; human cell line; porcine cell line; ss.
 OS Pig endogenous retrovirus.

XX Key Location/Qualifiers
 XX CDS 211..2193
 XX /tag= a
 XX /product= envelope_protein

PN M09853104-A2.

PD 26-NOV-1998.

PF 18-MAY-1998; 98MO-GB01428.

PR 16-MAY-1997; 97GB-0010154.

PA (MED1-) MEDICAL RES COUNCIL.

PI Stoye JP, Weiss RA;

PI WPI: 1999-045324/04.

DR P-PSDB; AAM85453.

PT Newly isolated nucleic acid probe capable of hybridizing to either
 PT the PERV-A or PERV-B env gene - useful in the detection of
 PT retroviruses, and their subtypes, in a sample of porcine/human
 PT tissue

PS Claim 6; Page 20-21; 36pp; English.

XX The present sequence encodes a pig endogenous retrovirus (PERV)-A
 CC envelope protein. PERV exists in two different subtypes, PERV-A and
 CC PERV-B. The differences are reflected in sequence divergence in the
 CC envelope genes. Probes and primers can be derived from the envelope
 CC (env) genes of PERV-A and PERV-B. The probes and primers are used in
 CC a method to detect retroviruses in a sample of porcine/human tissue,
 CC particularly primary porcine tissue and human cell lines that have been
 CC cultivated in the presence of a porcine cell line, or human tissue from
 CC a patient with a xenotransplant. Subtype of PERV in a sample containing
 CC one of the PERV env genes can also be determined.

SO Sequence 2462 BP; 715 A; 573 C; 575 G; 599 T; 0 other;

Query Match 23.7%; Score 786; DB 20; Length 2462;

Best Local Similarity 99.9%; Pred. No. 1,1e-309; Indels 0; Gaps 0;

Matches 836; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2459 tcgaatgggtgtagcagcagcgtggaagcagctccggagagcctactcaaggagagact 2518
 DB 1 ccgagtggtgtagcagcagcgtggaagcagctccggagagcctactcaaggagagact 60
 OY 2519 tggcaagttccacatcgcttccaagttgagagttcagttatgtagagcgcacgtgag 2578
 DB 61 tggcaagttccacatcgcttccaagttgagagttcagttatgtagagcgcacgtgag 120
 OY 2579 gaaacctgagagctcggtggaagggacattctcgtactttgaccacacacagcgtg 2638
 DB 121 gaaacctgagagctcggtggaagggacattctcgtactttgaccacacacagcgtg 180
 OY 2639 tgaagtcgaaggaatcccccacttgatcctatgcatcccaagcttaagcgcgcacactc 2698
 DB 181 tgaagtcgaaggaatcccccacttgatcctatgcatcccaagcttaagcgcgcacactc 240
 OY 2699 ccgattcgggttggaagccggaagactggaatcccccacttaagcttgcctcactgcgc 2758
 DB 241 ccgattcgggttggaagccggaagactggaatcccccacttaagcttgcctcactgcgc 300
 OY 2759 tggctcctaactctcctaactcctaagcttaagcttaagcgcgcacactc 2818
 DB 301 tggctcctaactctcctaactcctaagcttaagcttaagcgcgcacactc 360

OY 2819 aactccataaaccttactctcactcactggttacttactgactcgcgtacaggtatcat 2878
 DB 361 aactccataaaccttactctcactcactggttacttactgactcgcgtacaggtatcat 420
 OY 2879 attacaagaactcaaggagagctcccttgggagcttgggtcgtgaattatagctgc 2938
 DB 421 attacaagaactcaaggagagctcccttgggagcttgggtcgtgaattatagctgc 480
 OY 2939 ctctgacagtaatcccttgggtcgtgaattatagctgc 2998
 DB 481 ctctgacagtaatcccttgggtcgtgaattatagctgc 540
 OY 2999 taagggtttagcttgggtcgtgaattatagctgc 3058
 DB 541 taagggtttagcttgggtcgtgaattatagctgc 600
 OY 3059 gattccttgggtcgtgaattatagctgc 3118
 DB 601 gattccttgggtcgtgaattatagctgc 660
 OY 3119 gttcctcagcagagagtagttagcttgggtcgtgaattatagctgc 3178
 DB 661 gttcctcagcagagagtagttagcttgggtcgtgaattatagctgc 720
 OY 3179 tttaattagcagatggagatggagatggagatggagatggagatggagatggagatggag 3238
 DB 721 tttaattagcagatggagatggagatggagatggagatggagatggagatggagatggag 780
 OY 3239 aataagaacaataagctgttactcgtttagacacttacttaaaaaaagttactc 3295
 DB 781 aataagaacaataagctgttactcgtttagacacttacttaaaaaaagttactc 837

RESULT 10

AC67032

ID AAC67032 standard; DNA; 6076 BP.

AC67032;

27-MAR-2001 (first entry)

PERV env protein coding sequence SEQ ID NO: 32.

Xenotransplantation; infectious agent; vaccine; ds.

Porcine endogenous retrovirus.

MO200071726-A1.

30-NOV-2000.

24-MAY-2000; 2000MO-US14296.

24-MAY-1999; 99US-0135631.

(MAYO-) MAYO MEDICAL VENTURES.

Federspiel MJ;

WPI: 2001-032041/04.

Inhibiting or preventing infectious agent transmission in mammalian

transplant recipients, by introducing recombinant DNA comprising DNA

encoding extracellular proteins of the agent into donor cells, such as

swine cells -

Claim 16; Page 117-119; 14pp; English.

The present invention provides a method to prevent the transmission of

infectious agents during xenotransplantation. This involves introducing

to donor swine cells a recombinant DNA encoding a peptide fragment from

the infectious agent, and then introducing these cells into the

CC transplant recipient.

XX Sequence 6076 BP, 1619 A; 1505 C; 1598 G; 1354 T; 0 other;

Query Match 22.5%; Score 746; DB 22; Length 6076;
Best Local Similarity 98.7%; Pred. No. 1,66-293;
Matches 2196; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 31 tctgtcctagatatacaccactagccacacttttctgcttctggaatgagagatcca 90
Db 3577 tctgtcctagatatacaccactagccacacttttctgcttctggaatgagagatcca 3636
QY 91 ggtacggagaagaccggcgagctcaccttgagccgagctgcccgaagggtcaagaatccc 150
Db 3637 ggtacggagaagaccggcgagctcaccttgagccgagctgcccgaagggtcaagaatccc 3696
QY 151 ccgacatctttagagaaagccctacacagggagcccgcccaactttagatccaacct 210
Db 3697 ccgacatctttagagaaagccctacacagggagcccgcccaactttagatccaacct 3756
QY 211 caggtgacccctcctcagtaagtgatatacctgtcttcgcgagggagccacaagaagac 270
Db 3757 caggtgacccctcctcagtaagtgatatacctgtcttcgcgagggagccacaagaagac 3816
QY 271 tggctagaagtgacgaagagcactactgtggaattgtctgaactagctacagaagccct 330
Db 3817 tggctagaagtgacgaagagcactactgtggaattgtctgaactagctacagaagccct 3876
QY 331 gctaaagaagcccgagatttgcagagagaaglaacatacttgggtggtacagtttcggggc 390
Db 3877 gctaaagaagcccgagatttgcagagagaaglaacatacttgggtggtacagtttcggggc 3936
QY 391 gggcagcgatgtgacgagagcgacggaagaanaactgagtcgagataccggcccacac 450
Db 3937 gggcagcgatgtgacgagagcgacggaagaanaactgagtcgagataccggcccacac 3996
QY 451 acagcgcaacaagaatgagagagtttttgggagcagctggaatttgcagagctgtgatalccg 510
Db 3997 acagcgcaacaagaatgagagagtttttgggagcagctggaatttgcagagctgtgatalccg 4056
QY 511 gggcttgagaccttagacggcccccctacccgctaacccaagaagaaggggatcttc 570
Db 4057 gggcttgagaccttagacggcccccctacccgctaacccaagaagaaggggatcttc 4116
QY 571 tgggtcctcctgagcaccagaagcatttgcgtctatacgaagaagccctgcgagcgacac 630
Db 4117 tgggtcctcctgagcaccagaagcatttgcgtctatacgaagaagccctgcgagcgacac 4176
QY 631 gctctgacctcctgacgtaactaaaccttaaccttatagttgaatgaagcgtlaaagg 690
Db 4177 gctctgacctcctgacgtaactaaaccttaaccttatagttgaatgaagcgtlaaagg 4236
QY 691 gtagccggagaggttttaaccacacaccttagacatagggagagacctgttgcctactcg 750
Db 4237 gtagccggagaggttttaaccacacaccttagacatagggagagacctgttgcctactcg 4296
QY 751 tcaagaagactgtatccctgtlagccaggtgttgcggctatgctcgaaggtatccgaagct 810
Db 4297 tcaagaagactgtatccctgtlagccaggtgttgcggctatgctcgaaggtatccgaagct 4356
QY 811 gtggtcctactgttcaagaagcgtgtgacaatgtgaacttgggagacagataataactgtta 870
Db 4357 gtggtcctactgttcaagaagcgtgtgacaatgtgaacttgggagacagataataactgtta 4416
QY 871 gccccccatgcatgtgagaacatcggttcggcagcccccaacggatgagatgagacacg 930
Db 4417 gccccccatgcatgtgagaacatcggttcggcagcccccaacggatgagatgagacacg 4476
QY 931 cgaatgacccatacgaagcgtgtcttcacagagaaggttcaacttgcctccacagc 990
Db 4477 cgaatgacccatacgaagcgtgtcttcacagagaaggttcaacttgcctccacagc 4536

QY 991 gctcacaacctgccaactcttctgtcgtgaagagatgatgaaccagtgactcatgttc 1050
Db 4537 gctcacaacctgccaactcttctgtcgtgaagagatgatgaaccagtgactcatgttc 4596
QY 1051 catcaactatgtatgaagagacgtggttcccgcaagagaccttaacagataccgtgact 1110
Db 4597 catcaactatgtatgaagagacgtggttcccgcaagagaccttaacagataccgtgact 4656
QY 1111 ggaagaatgtcaacctgttctactgacgggaagcagctatgttggtagaaggtlaagagatg 1170
Db 4657 ggaagaatgttcaacctgttctactgacgggaagcagctatgttggtagaaggtlaagagatg 4716
QY 1171 gctggggcgagctggtgagacgggagccgacagatctcgggcccagccgtcgcggaaga 1230
Db 4717 gctggggcgagctggtgagacgggagccgacagatctcgggcccagccgtcgcggaaga 4776
QY 1231 acttcaagcgcaaaagcgtgagcttcaatgacctcagcgaagacttgcggtcgcggaagg 1290
Db 4777 acttcaagcgcaaaagcgtgagcttcaatgacctcagcgaagacttgcggtcgcggaagg 4836
QY 1291 aaatccataacattatacaggaacagagatagcttgcgactgcacagctacacgg 1350
Db 4837 aaatccataacattatacaggaacagagatagcttgcgactgcacagctacacgg 4896
QY 1351 gcaatctataacaaagggttgcctactcagcgagggggaagaaataaagaacaaagg 1410
Db 4897 gcaatctataacaaagggttgcctactcagcgagggggaagaaataaagaacaaagg 4956
QY 1411 gaaattctaaagcctatagaagccttaacatttgcgaagaagctatataactgt 1470
Db 4957 gaaattctaaagcctatagaagccttaacatttgcgaagaagctatataactgt 5016
QY 1471 cctggagataagaagaagccaaagatcctacatactagaagggaacagatgctgcagggt 1530
Db 5017 cctggagataagaagaagccaaagatcctacatactagaagggaacagatgctgcagggt 5076
QY 1531 gccaaagcagcagcccgagctgtttaaactcttcctatataagaacggcccgaaccca 1590
Db 5077 gccaaagcagcagcccgagctgtttaaactcttcctatataagaacggcccgaaccca 5136
QY 1591 gaaaccaagcagcagtaacacctagaagactggcgaagaagataaanaagataagccagttc 1650
Db 5137 gaaaccaagcagcagtaacacctagaagactggcgaagaagataaanaagataagccagttc 5196
QY 1651 tctgaagactccggaaggagcctgtctactactatataaggaaggaatcctgcgcccaaa 1710
Db 5197 tctgaagactccggaaggagcctgtctactactatataaggaaggaatcctgcgcccaaa 5256
QY 1711 gaaagggtagaatatgtccaacagatatacgtcttaacccacctagaagactaaacacgtg 1770
Db 5257 gaaagggtagaatatgtccaacagatatacgtcttaacccacctagaagactaaacacgtg 5316
QY 1771 cagcagttgttcaagaacatcccttcaatcagtcttgcaggtacacaggaatgtgctactcg 1830
Db 5317 cagcagttgttcaagaacatcccttcaatcagtcttgcaggtacacaggaatgtgctactcg 5376
QY 1831 gtgtgtaaacatgtgtgtgcctcgccaagctgtgttaatgtcaatccttccaagaatccoca 1890
Db 5377 gtgtgtaaacatgtgtgtgcctcgccaagctgtgttaatgtcaatccttccaagaatccoca 5436
QY 1891 ggaagaagactaaggggaagccacacagcgctcaactggaagtgagacttaagta 1950
Db 5437 ggaagaagactaaggggaagccacacagcgctcaactggaagtgagacttaagta 5496
QY 1951 aagccggcctaaatacggaaacataatctatgtgttcttgaagacactttcaagatg 2010
Db 5497 aagccggcctaaatacggaaacataatctatgtgttcttgaagacactttcaagatg 5556
QY 2011 gtagaggtctacccctaaagaagaagacttcaacccgtgtgtgtgtgtgtgtgtgtgtgtgt 2070
Db 5557 gtagaggtctacccctaaagaagaagacttcaacccgtgtgtgtgtgtgtgtgtgtgtgtgt 5616
QY 2071 gaaattttccaagaatttgaatacctaaagtaagtaaggtcagacatgtctcccttc 2130

```

DB 5617 gaattcttcccaagatttggatacttaagcgaatggttcgacacatggtccagcttc 5676
QY 2131 gttcccaagtaagtaaggaagactggtcccaagatattggagattgattggaactgattgt 2190
DB 5677 gttcccaagtaagtaaggaagactggtcccaagatattggagattgattggaactgattgt 5736
QY 2191 gctacagaccccaagcctcagacaggtagagagattgattgaaacattaaagacc 2250
DB 5737 gctacagaccccaagcctcagacaggtagagagattgattgaaacattaaagacc 5796
QY 2251 cttac 2255
DB 5797 cttac 5801

RESULT 11
AAT74884
ID AAT74884 standard; cDNA; 7892 BP.
XX
AC AAT74884;
XX
DT 09-FEB-1998 (first entry)
XX
DE Miniature swine retrovirus cDNA.
XX
KW Retrovirus; porcine; GAG protein; POL protein; ENV protein;
KW xenotransplantation; infectious; provirus; organ transplant; donor;
KW activated virus; PCR; ss.
XX
OS Porcine retrovirus.
XX
FH Key Location/Qualifiers
FT CDS 585..2159
FT /tag= a
FT mat-peptide 585..2156
FT /tag= b
FT /note= "putative GAG protein"
FT CDS 2307..5744
FT /tag= c
FT mat-peptide 2307..5741
FT /tag= d
FT /note= "putative POL protein"
FT CDS 5620..7536
FT /tag= e
FT mat-peptide 5620..7533
FT /tag= f
FT /note= "putative ENV protein"
XX
PN WO9721836-A1.
XX
PD 19-JUN-1997.
XX
PE 13-DEC-1996; 96WO-US19680.
XX
PR 14-DEC-1995; 95US-0572645.
XX
PA (GENO ) GEN HOSPITAL CORP.
XX
PI Fishman JA;
XX
DR WPI; 1997-332804/30.
DR P-PSDB; AAW32096-W32098.
XX
PT New nucleic acid from porcine retrovirus - used for detecting
PT viruses in transplant or other tissue and for assessing risk of
PT transmitting infection to graft recipient
XX
XX Claim 22; Fig 3; 128bp; English.
XX
CC This cDNA sequence represents a porcine retrovirus from miniature swine
CC containing the putative coding regions for viral GAG, POL and ENV
CC proteins. This sequence and PCR fragments generated from it

```

```

CC (see AAT74812-T74882) can be used to screen organs for the presence of
CC porcine retroviruses prior to xenotransplantation. Transplantation can
CC increase the likelihood of retroviral activation if intact and
CC infectious proviruses are present. The porcine retroviral sequence can be
CC used to generate probes to determine the level (e.g. copy number) of
CC intact (i.e. potentially replicating) porcine provirus sequences in a
CC strain of xenograft transplantation donors. It can be used to detect
CC mutations, genetic lesions or viral recombinants and also to determine
CC the histological localization of activated retroviruses. Using Polymerase
CC Chain Reaction DNA Quantitation (PQD) on blood mononuclear cells,
CC infectivity titration and susceptibility testing can be performed.
CC Ultimately animal donors without intact porcine retroviral sequences or a
CC lower copy number of viral elements could be selected.
XX
SQ Sequence 7892 BP; 2191 A; 1915 C; 1980 G; 1806 T; 0 other;

Query Match 21.2%; Score 704; DB 18; Length 7892;
Best Local Similarity 98.7%; Pred. No. 1.7e-276;
Matches 2204; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 22 gatgcctctctctgctgagattaccccaactagcccaacttttgcttgaatg 81
DB 2970 gatgcctctctctgctgagattaccccaactagcccaacttttgcttgaatg 3029
QY 82 agagatccaggtacggaagaacggtgagctcactgacgcagctgcccgaaggttc 141
DB 3030 agagatccaggtacggaagaacggtgagctcactgacgcagctgcccgaaggttc 3089
QY 142 aagaactcccgacacatcttgagaaagccctacacagagccctggaacttcagatc 201
DB 3090 aagaactcccgacacatcttgagaaagccctacacagagccctggaacttcagatc 3149
QY 202 caacacccctcagtgacctctctcagatcgtgtgacgtctctggtcggtgacac 261
DB 3150 caacacccctcagtgacctctctcagatcgtgtgacgtctctggtcggtgacac 3209
QY 262 aaacaggaactgcttgaaggtacgaagcactacgtcgtgaattgtctgacctggctac 321
DB 3210 aaacaggaactgcttgaaggtacgaagcactacgtcgtgaattgtctgacctggctac 3269
QY 322 agagcctctgctaaagaagcccaagatttgcaagagaaggttaacatcttggtgacagt 381
DB 3270 agagcctctgctaaagaagcccaagatttgcaagagaaggttaacatcttggtgacagt 3329
QY 382 ttgcggtggtggtgagatggtgctgaacggaagcaagaactgtatgtccagatccg 441
DB 3330 ttgcggtggtggtgagatggtgctgaacggaagcaagaactgtatgtccagatccg 3389
QY 442 gccccaacacagccaagaatgagagatgttttggtggaacgtggaatttgagagctg 501
DB 3390 gccccaacacagccaagaatgagagatgttttggtggaacgtggaatttgagagctg 3449
QY 502 tgaatcccggtgttgagacacttaacgcccactctacccgcttaacccaagaagaagg 561
DB 3450 tgaatcccggtgttgagacacttaacgcccactctacccgcttaacccaagaagaagg 3509
QY 562 ggaattctctggtgctctgagcacaagaagcatttgatgctatacaaaaagccctgctg 621
DB 3510 ggaattctctggtgctctgagcacaagaagcatttgatgctatacaaaaagccctgctg 3569
QY 622 aggcacactgctctggtgctctgagcgaactaacccttaaccccttatgtgagatgag 681
DB 3570 aggcacactgctctggtgctctgagcgaactaacccttaaccccttatgtgagatgag 3629
QY 682 cgtaaagagtagcccggaagaglttaacccaacacttagacacatggaagagacgtgt 741
DB 3630 cgtaaagagtagcccggaagaglttaacccaacacttagacacatggaagagacgtgt 3689
QY 742 gctactgtcaagaagacttactctgtagcaggtgtgtggtccgtatgttgaagct 801
DB 3690 gctactgtcaagaagacttactctgtagcaggtgtgtggtccgtatgttgaagct 3749

```

QY 802 atcgcaactgtgacatactgctcaagagcgtgcacaatlgaacttgggacagaata 861
|||||
Db 3750 atcgcaactgtgacatactgctcaagagcgtgcacaatlgaacttgggacagaata 3809
QY 862 acttaataagcccccatgcatctggaagaacatgcttcggaagccccagacgatgata 921
|||||
Db 3810 actgtaataagcccccatgcatctggaagaacatgcttcggaagccccagacgatgata 3869
QY 922 accaagcccgcaatgaccccaatcaaaagcctgtctctcaagaagggctcaatttcgct 981
|||||
Db 3870 accaagcccgcaatgaccccaatcaaaagcctgtctctcaagaagggctcaatttcgct 3929
QY 982 ccaacccgacgctctcaacccctgcactctctgctcgtgaagagactgtgaacacgtact 1041
|||||
Db 3930 ccaacccgacgctctcaacccctgcactctctgctcgtgaagagactgtgaacacgtact 3989
QY 1042 catgattgcatcaactatgtattgtgagagactggggtccgcaaggaacttcaagacata 1101
|||||
Db 3990 catgattgcatcaactatgtattgtgagagactggggtccgcaaggaacttcaagacata 4049
QY 1102 ccgctgactggaagaatgttaaacctgttcaactgcgaagacgactatgtgtgaaagt 1161
|||||
Db 4050 ccgctgactggaagaatgttaaacctgttcaactgcgaagacgactatgtgtgaaagt 4109
QY 1162 aagagagatgtgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1221
|||||
Db 4110 aagagagatgtgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 4169
QY 1222 ccggaagagacttcaacgcaaaagcgtgactgactgactgactgactgactgactgactg 1281
|||||
Db 4170 ccggaagagacttcaacgcaaaagcgtgactgactgactgactgactgactgactgactg 4229
QY 1282 gccgaaggggaatcctaataacattatacgaagcaggtatgacttctgcaactgacac 1341
|||||
Db 4230 gccgaaggggaatcctaataacattatacgaagcaggtatgacttctgcaactgacac 4289
QY 1342 gfacacggggccatctataaaaggggtgtgttacttactcgaagggaggaataaag 1401
|||||
Db 4290 gfacacggggccatctataaaaggggtgtgttacttactcgaagggaggaataaag 4349
QY 1402 aacaaaggggaatcctaataacattatacgaagcaggtatgacttctgcaactgacac 1461
|||||
Db 4350 aacaaaggggaatcctaataacattatacgaagcaggtatgacttctgcaactgacac 4409
QY 1462 ataactgtctgtgacatcagaagaacaaatcatalatagagggaaacagatgct 1521
|||||
Db 4410 ataactgtctgtgacatcagaagaacaaatcatalatagagggaaacagatgct 4469
QY 1522 gaccgggttgcgaagcagcagccagctgttaaccttgcctataatagaacgccc 1581
|||||
Db 4470 gaccgggttgcgaagcagcagccagctgttaaccttgcctataatagaacgccc 4529
QY 1582 aaagcccaagaacccaagcagcagctacacccctagaagaactgcaagagataaaagata 1641
|||||
Db 4530 aaagcccaagaacccaagcagcagctacacccctagaagaactgcaagagataaaagata 4589
QY 1642 gaccagttcttctgagactccggaaggagcctgtatatacctcaatgtggaagaaactct 1701
|||||
Db 4590 gaccagttcttctgagactccggaaggagcctgtatatacctcaatgtggaagaaactct 4649
QY 1702 ccccaacaagaaggttagaatatgtcaacagatactgttcaaacccactgagact 1761
|||||
Db 4650 ccccaacaagaaggttagaatatgtcaacagatactgttcaaacccactgagact 4709
QY 1762 aaacaactgcagcagctgtgtagaacatcccttatcatgttctgaggtaccagagatg 1821
|||||
Db 4710 aaacaactgcagcagctgtgtagaacatcccttatcatgttctgaggtaccagagatg 4769
QY 1822 gctgactcgtgtgcaaacattgtgtccctgcagcagctgtgttaattgttaactcctcaga 1881
|||||
Db 4770 gctgactcgtgtgcaaacattgtgtccctgcagcagctgtgttaattgttaactcctcaga 4829
QY 1882 atactccaggaagaagactlaagggaagcaccagcgctcaactggtggaagtgcacttc 1941

Db 4830 atgctcccaagggaagaagactlaagggaagccacccagcgctcactggtgaagtgcacttc 4889
|||||
QY 1942 actgaggttaaaagccggtctaataatcaggaacaaatactatgtgtttttgtagacactt 2001
|||||
Db 4890 actgaggttaaaagccggtctaataatcaggaacaaatactatgtgtttttgtagacactt 4949
QY 2002 tcaagatgtgtgaggttatctacttaagaagaagacttcaacccgtgtgtgtaagaa 2061
|||||
Db 4950 tcaagatgtgtgaggttatctacttaagaagaagacttcaacccgtgtgtgtaagaa 5009
QY 2062 atactggaaggaatttcccaagatttgaataccttaaggttaaggttcaagaaatgt 2121
|||||
Db 5010 atactggaaggaatttcccaagatttgaataccttaaggttaaggttcaagaaatgt 5069
QY 2122 ccagcttctgtgtcccaaggttaagtcaggtggtcccaagatatgtggtattgtgaa 2181
|||||
Db 5070 ccagcttctgtgtcccaaggttaagtcaggtggtcccaagatatgtggtattgtgaa 5129
QY 2182 ctgcatgtgtatagaagcccaaaagctcaggaacaggttaagagagataagaacact 2241
|||||
Db 5130 ctgcatgtgtatagaagcccaaaagctcaggaacaggttaagagagataagaacact 5189
QY 2242 aaagagacccttac 2255
|||||
Db 5190 aaagagacccttac 5203

RESULT 12

AAF77727 standard; DNA: 8132 BP.

AAF77727:

23-MAY-2001 (first entry)

Nucleotide sequence of a retrovirus found in miniature swine.

Retrovirus; graft transplantation; xenotransplantation; miniature swine; ss.

Unidentified.

US6190861-B1.

20-FEB-2001.

13-DEC-1996; 9605-0766528.

14-DEC-1995; 9505-0572645.

(GENO) GEN HOSPITAL CORP.

Fishman JA;

WPI: 2001-256211/26.

P-P-SDB; AAB73285, AAB73286, AAB73287.

Assessing risk of endogenous retroviruses in clinical practice and in xenotransplantation, comprises using probe sequences derived from swine or miniature swine retroviral genome -

Claim 1, Fig 3; 127pp; English.

The present invention relates to a method for screening a cell or tissue for the presence or expression of a retrovirus (RV), comprising contacting a target nucleic acid from the cell or tissue with a second nucleic acid (a fragment thereof). The method is useful for RV detection and to assess graft transplantation risk. Screening of animals allows the elimination of donors with active replication of known viruses. Inactive proviruses can be detected and inactivated, allowing identification and elimination of potential human pathogens derived from swine in a manner not possible

QY 2062 atactgaggaatttttccaagatttgaataactaagtaataggttcagacaatgt 2121
|||||
DB 5010 atactggaagaatttttccaagatttgaataactaagtaataggttcagacaatgt 5069
QY 2122 ccagcttctgtccaggaagtaagtcagagactgcccagaatactgggagatgtaagaa 2181
|||||
DB 5070 ccagcttctgtccaggaagtaagtcagagactgcccagaatactgggagatgtaagaa 5129
QY 2182 ctgcatctgcatatcacagaccaccaagctcagagacagtagagagagatgaataaccat 2241
|||||
DB 5130 ctgcatctgcatatcacagaccaccaagctcagagacagtagagagagatgaataaccat 5189
QY 2242 aaagagacccttac 2255
|||||
DB 5190 aaagagacccttac 5203

RESULT 13
AAC67022
ID AAC67022 standard; DNA; 4918 BP.
XX
AC AAC67022;
XX
DT 27-MAR-2001 (first entry)
XX
DE PERV env protein coding sequence SEQ ID NO: 22.
XX
XX Xenotransplantation; infectious agent; vaccine; ds.
XX
OS Porcine endogenous retrovirus.
XX
PN WO200071726-A1.
XX
PD 30-NOV-2000.
XX
PF 24-MAY-2000; 2000WO-US14296.
XX
PR 24-MAY-1999; 99US-0135631.
XX
PA (MAYO-) MAYO MEDICAL VENTURES.
XX
PI Federspiel MJ;
XX
XX WPI; 2001-032041/04.
DR
XX Inhibiting or preventing infectious agent transmission in mammalian
PT transplant recipients, by introducing recombinant DNA comprising DNA
PT encoding extracellular proteins of the agent into donor cells, such as
PT swine cells -
XX
PS Claim 16; Page 109-111; 144pp; English.
XX
CC The present invention provides a method to prevent the transmission of
CC infectious agents during xenotransplantation. This involves introducing
CC to donor swine cells a recombinant DNA encoding a peptide fragment from
CC the infectious agent, and then introducing these cells into the
CC transplant recipient.
XX
SQ Sequence 4918 BP; 1382 A; 1173 C; 1183 G; 1180 T; 0 other;

Query Match 17.2%; Score 570; DB 22; Length 4918;
Best Local Similarity 98.7%; Pred. No. 4.4e-222;
Matches 1670; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 684 taaggagtagcccgaggaggttttaaccaccaaccttagggacatggagagactgtg 743
|||||
DB 199 taaggagtagcccgaggaggttttaaccaccaaccttagggacatggagagactgtg 258
QY 744 ctactgtccaagaagactgtgactgtgactgtgactgtgactgtgactgtgactgtgactgtg 803
|||||
DB 259 ctactgtccaagaagactgtgactgtgactgtgactgtgactgtgactgtgactgtgactgtg 318
QY 804 cgcagctgtgagcactgtgactgtgactgtgactgtgactgtgactgtgactgtgactgtg 863
|||||
DB 319 cgcagctgtgagcactgtgactgtgactgtgactgtgactgtgactgtgactgtgactgtg 378
QY 864 tgtatagcccccatgcatgtgaggaacatgtgtccgagcccccagaccagatgtgac 923
|||||
DB 379 tgtatagcccccatgcatgtgaggaacatgtgtccgagcccccagaccagatgtgac 438
QY 924 caacgcccgaatgacccactatcaaaagctgtgttccacagagaggttcaacttgcctcc 983
|||||
DB 439 caacgcccgaatgacccactatcaaaagctgtgttccacagagaggttcaacttgcctcc 498
QY 984 accagccgcttcaaccctgacactcttctcctggaagagatgataaccagatgtacta 1043
|||||
DB 499 accagccgcttcaaccctgacactcttctcctggaagagatgataaccagatgtacta 558
QY 1044 tgattgcatcaactatgtatgtgagagactgggtccgcaagagaccttacaagacatacc 1103
|||||
DB 559 tgattgcatcaactatgtatgtgagagactgggtccgcaagagaccttacaagacatacc 618
QY 1104 gctgactgagagagtgtaactgttcaactgtgactgtgactgtgactgtgactgtgactgtg 1163
|||||
DB 619 gctgactgagagagtgtaactgttcaactgtgactgtgactgtgactgtgactgtgactgtg 678
QY 1164 gaggaatgt 1223
|||||
DB 679 gaggaatgt 738
QY 1224 ggaaggaacttcagcgaagaagctgtgactgtgactgtgactgtgactgtgactgtgactgtg 1283
|||||
DB 739 ggaaggaacttcagcgaagaagctgtgactgtgactgtgactgtgactgtgactgtgactgtg 798
QY 1284 cgaaggaacttcataaactattatgaagcagcagatgtgactgtgactgtgactgtgactgtg 1343
|||||
DB 799 cgaaggaacttcataaactattatgaagcagcagatgtgactgtgactgtgactgtgactgtg 858
QY 1344 acacggggccatctataaacaagggt 1403
|||||
DB 859 acacggggccatctataaacaagggt 918
QY 1404 caaagagaatcttcaagcctatgaagccttcaactgtgcaaaaaggctagtattat 1463
|||||
DB 919 caaagagaatcttcaagcctatgaagccttcaactgtgcaaaaaggctagtattat 978
QY 1464 acactgtcttgatcatcagaagaagcccaagatcctatctgaggaagccagatgtgctga 1523
|||||
DB 979 acactgtcttgatcatcagaagaagcccaagatcctatctgaggaagccagatgtgctga 1038
QY 1524 ccgggttccgaagcagcagccagcgtgttaaccttgcctataatagaagacgcccac 1583
|||||
DB 1039 ccgggttccgaagcagcagccagcgtgttaaccttgcctataatagaagacgcccac 1098
QY 1584 agccccaagaaccagacagatcacacctagaagactgtgcaagagataaagaataga 1643
|||||
DB 1099 agccccaagaaccagacagatcacacctagaagactgtgcaagagataaagaataga 1158
QY 1644 ccagttctctgagactccggagggagactgtctatccctcatatggaaggaatctctgccc 1703
|||||
DB 1159 ccagttctctgagactccggagggagactgtctatccctcatatggaaggaatctctgccc 1218
QY 1704 ccacaagaaggttagaatatgtccacaagaataatgtcttaaccacccatagaactaa 1763
|||||
DB 1219 ccacaagaaggttagaatatgtccacaagaataatgtcttaaccacccatagaactaa 1278
QY 1764 acactgtcagcagttgtgtcagaacatcccttatcatgttctgaggtaccagagatggc 1823

DB 1279 acacccgagcaggttgcgaacatctctctatcatgttctgagcaccagagatgac 1338
QY 1824 tgaactcgtgtgcgaacatctgtccctcagctggttaagtaacccctcagat 1883
DB 1339 tgaactcgtgtgcgaacatctgtccctcagctggttaagtaacccctcagat 1398
QY 1884 accctcag 1943
DB 1399 accctcag 1458
QY 1944 tgaagtaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2003
DB 1459 tgaagtaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1518
QY 2004 aggaatggttag 2063
DB 1519 aggaatggttag 1578
QY 2064 actgag 2123
DB 1579 actgag 1638
QY 2124 agcttcgtccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2183
DB 1639 agcttcgtccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1698
QY 2184 gcatgtgcatacag 2243
DB 1699 gcatgtgcatacag 1758
QY 2244 agagacccctac 2255
DB 1759 agagacccctac 1770

RESULT 14

AAH26252
ID AAH26252 standard; DNA; 600 BP.

AC AAH26252;

DT 02-OCT-2001 (first entry)

DE Pig endogenous retrovirus clone 12002-2.

KM PERV; pig; xenotransplantation; ss.

OS Pig endogenous retrovirus.

PN WO200153825-A1.

XX 26-JUL-2001.

PF 19-JAN-2001; 2001MO-US01857.

XX 19-JAN-2000; 2000US-0177003.

PR 16-FEB-2000; 2000US-0182965.

PR 27-OCT-2000; 2000US-0243695.

XX (BIOT-) BIOTRANSPLANT INC.

XX Patience C;

XX WPI; 2001-483157/52.

XX New inbred swine of DD haplotype defective for transmission of porcine

XX endogenous retroviruses, useful as source of clinical organs, tissues

XX and cells, or as source of stem cells for xenogenic stem cell and

XX thymic replacement therapy -

XX Example 4; Fig 3; 66pp; English.

CC The present sequence is that of pig endogenous retrovirus (PERV)
CC PCR-amplified clone 12002-2. Peripheral blood mononuclear cells
CC from swine leukocyte antigen inbred miniature swine number 12002
CC or 11619 were induced for PERV expression, and then co-cultured
CC with uninfected human 293 cells. The infection of the cells was
CC determined by the presence of reverse transcriptase (RT) activity
CC in the culture supernatants. Genomic DNA was isolated and PCR
CC products were generated using LTR-to-LTR PCR from cultures showing
CC high levels of RT activity. The PCR products were cloned and
CC sequenced to determine the different types of replication competent
CC PERV present. 7 PERV sequences were derived from pig 12002 (see
CC AAH26251-57) and 9 from pig 11619 (see AAH26258-66). Having
CC determined the spectrum of sequences capable of replicating in
CC human cells it is then possible to determine if these sequences are
CC present in the genome of D/D haplotype miniature swine. Their
CC absence then shows a lack of production of human-tropic virus from
CC the cells of the D/D haplotype animals. The object of the invention
CC is to provide a swine breed which does not produce PERV capable of
CC infecting human cells, as measured by high-stringency in vitro
CC assays. Such animals, which are preferably miniature pigs of the
CC DD haplotype, have a unique advantage over other swine breeds with
CC respect to microbiological safety for xenotransplantation. They
CC are a valuable source of organs, tissues and cells, including a
CC source of stem cells for xenogenic stem cell and thymic replacement
CC therapy, e.g. in treating AIDS cases. Methods of producing such
CC animals and for screening animals for the desired properties are
CC disclosed.

SO Sequence 600 BP; 166 A; 149 C; 129 G; 156 T; 0 other;

Query Match 15.0%; Score 498; DB 22; Length 600;
Best Local Similarity 99.7%; Pred. No. 1e-192;
Matches 598; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2663 atgcatcccaagtttaagccgagcagcagcagcagcagcagcagcagcagcagcagc 2728
DB 1 atgcatcccaagtttaagccgagcagcagcagcagcagcagcagcagcagcagcagc 60
QY 2729 aaatcccttaagcttgcgcagcagcagcagcagcagcagcagcagcagcagcagcagc 2788
DB 61 aaatcccttaagcttgcgcagcagcagcagcagcagcagcagcagcagcagcagcagc 120
QY 2789 gtaatgtaaaagccttcttgacagcagcagcagcagcagcagcagcagcagcagcagc 2848
DB 121 gtaatgtaaaagccttcttgacagcagcagcagcagcagcagcagcagcagcagcagc 180
QY 2849 ttaactactgactccggtgacaggtatataatcaacagcagcagcagcagcagcagcagc 2908
DB 181 ttaactactgactccggtgacaggtatataatcaacagcagcagcagcagcagcagcagc 240
QY 2909 gggacctggtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2968
DB 241 gggacctggtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 300
QY 2969 cagggcacaaccccccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3028
DB 301 cagggcacaaccccccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 360
QY 3029 aataatgaagaatatttggaatctcaggaattcttcttcaagcaatgagcagcagcagc 3088
DB 361 aataatgaagaatatttggaatctcaggaattcttcttcaagcaatgagcagcagcagc 420
QY 3089 acttctaagtggaatatttggaatctcaggaattcttcttcaagcaatgagcagcagcagc 3148
DB 421 acttctaagtggaatatttggaatctcaggaattcttcttcaagcaatgagcagcagcagc 480
QY 3149 ttgttacaactctaccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3208
DB 481 ttgttacaactctaccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 540
QY 3209 tggcacaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3268

DB 541 tggcaacgcyggtacaaaagatgacgaataaagcaataaagctgcatcgttagac 600
RESULT 15
AAH26258
ID AAH26258 standard; DNA; 599 BP.
XX
XX AAH26258;
AC
XX
XX 02-OCT-2001 (first entry)
DT
XX
XX Pig endogenous retrovirus clone 11619-1.
DE
XX
XX PERV; pig; xenotransplantation; ss.
KM
XX
XX Pig endogenous retrovirus.
OS
XX
XX WO200153825-A1.
PN
XX
XX 26-JUL-2001.
PD
XX
XX 19-JAN-2001; 2001WO-US01857.
PE
XX
XX 19-JAN-2000; 2000US-0177003.
PR 16-FEB-2000; 2000US-0182965.
PR 27-OCT-2000; 2000US-0243695.
XX
XX (BIOT-) BIOTRANSPLANT INC.
PA
XX
XX Patience C;
PI
XX
XX WPI; 2001-483157/52.
DR
XX
XX
XX New inbred swine of DD haplotype defective for transmission of porcine
PT endogenous retroviruses, useful as source of clinical organs, tissues
PT and cells, or as source of stem cells for xenogenic stem cell and
PT thymic replacement therapy
PT
XX
XX Example 4; Fig 5; 66pp; English.
PS
XX
XX The present sequence is that of pig endogenous retrovirus (PERV)
CC PCR-amplified clone 11619-1. Peripheral blood mononuclear cells
CC from swine leukocyte antigen inbred miniature swine number 11619
CC or 12002 were induced for PERV expression, and then co-cultured
CC with uninfected human 293 cells. The infection of the cells was
CC determined by the presence of reverse transcriptase (RT) activity
CC in the culture supernatants. Genomic DNA was isolated and PCR
CC products were generated using LFR-to-LFR PCR from cultures showing
CC high levels of RT activity. The PCR products were cloned and
CC sequenced to determine the different types of replication competent
CC PERV present. 9 PERV sequences were derived from pig 11619 (see
CC AAH26258-66) and 7 from pig 12002 (see AAH26251-57). Having
CC determined the spectrum of sequences capable of replicating in
CC human cells it is then possible to determine if these sequences are
CC present in the genome of D/D haplotype miniature swine. Their
CC absence then shows a lack of production of human-tropic virus from
CC the cells of the D/D haplotype animals. The object of the invention
CC is to provide a swine breed which does not produce PERV capable of
CC infecting human cells, as measured by high-stringency in vitro
CC assays. Such animals, which are preferably miniature pigs of the
CC DD haplotype, have a unique advantage over other swine breeds with
CC respect to microbiological safety for xenotransplantation. They
CC are a valuable source of organs, tissues and cells, including a
CC source of stem cells for xenogenic stem cell and thymic replacement
CC therapy, e.g. in treating AIDS cases. Methods of producing such
CC animals and for screening animals for the desired properties are
CC disclosed.
XX
XX Sequence 599 BP; 166 A; 148 C; 129 G; 156 T; 0 other;

Query Match 15.0%; Score 497; DB 22; Length 599;
Best Local Similarity 99.7%; Pred. No. 2.6e-192;

	Matches	597;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps	0;
QY	2669	atgcatcccaacgttaagcgcgcacacccatccgattcgaggtygnaagccgnaaagactg								2728
DB	1	atgcatcccaacgttaagcgcgcacacccatccgattcgaggtygnaagccgnaaagactg								60
QY	2729	aaatcccttaagcttcgcctccatccgctggttcttactctgtcacaataactctcaa								2788
DB	61	aaatcccttaagcttcgcctccatccgctggttcttactctgtcacaataactctcaa								120
QY	2789	gttaatggttaaacgcttctggaacagccgaactccataaaccctatctcactctg								2848
DB	121	gttaatggttaaacgcttctggaacagccgaactccataaaccctatctcactctg								180
QY	2849	ttactactgactccggttaaggtatataataaacagcaactaaaggaggagctccctg								2908
DB	181	ttactactgactccggttaaggtatataataaacagcaactaaaggaggagctccctg								240
QY	2909	ggagcctggtggtcctgaatataatgtctgctcgttcgaactaaccctgtgtcgaatgac								2968
DB	241	ggagcctggtggtcctgaatataatgtctgctcgttcgaactaaccctgtgtcgaatgac								300
QY	2969	caagccacaccccccgatgactcgtgcttaacggttttaacglttgccaggaaccccca								3028
DB	301	caagccacaccccccgatgactcgtgcttaacggttttaacglttgccaggaaccccca								360
QY	3029	aataatgaagaataatgtggaatctcctcagagattccttgcgaagcaatgagctgcata								3088
DB	361	aataatgaagaataatgtggaatctcctcagagattccttgcgaagcaatgagctgcata								420
QY	3089	actctcaatgattggaatggaatggaatggaatggaatggaatggaatggaatggaatggaat								3148
DB	421	actctcaatgattggaatggaatggaatggaatggaatggaatggaatggaatggaatggaat								480
QY	3149	tttgttaacaatctcactacgaattataatcaatttaattatggtcgaatggaatggaatggaat								3208
DB	481	tttgttaacaatctcactacgaattataatcaatttaattatggtcgaatggaatggaatggaat								540
QY	3209	tggcaacgcyggtacaaaagatgacgaataaagcaataaagctgtcattcgttaga								3267
DB	541	tggcaacgcyggtacaaaagatgacgaataaagcaataaagctgtcattcgttaga								599

Search completed: February 24, 2002, 07:34:23
Job time: 16368 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2002, 07:42:18 ; Search time 584.4 Seconds

(without alignments)
4336.505 Million cell updates/sec

Title: US-09-171-553b-9

Perfect score: 2956

Sequence: 1 tgcctttagagtagaac.....aaaaaaaaaaaaaaaaa 2956

Scoring table: OLIGO NUC

Searched: 930621 seqs, 428662619 residues

Word size: 0

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

N_Geneseq_1101:*

- 1: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2191	74.1	8209	18	AAV09700 Porcine retrovirus
2	2089	70.7	8196	18	AAV09699 Porcine retrovirus
3	1872	63.3	1974	18	AAV09703 Porcine retrovirus
4	1743	59.0	3482	20	AAV82749 Pig endogenous ret
5	813	27.5	7333	22	AAV77726 Defective retrovir
6	457	15.5	2462	20	AAV82748 Pig endogenous ret
7	422	14.3	7393	18	AAV74883 Porcine retrovirus
8	143	4.8	3320	18	AAV09698 Porcine retrovirus
9	98	3.3	8132	22	AAV77727 Nucleotide sequenc
10	91	3.1	633	22	AAV67007 PERV-MSL LTR SEQ I
11	91	3.1	704	22	AAV67006 IamAI LTR SEQ ID N

12	91	3.1	4918	22	AAV67022 PERV env protein c
13	91	3.1	6076	22	AAV67021 PERV env protein c
14	91	3.1	6076	22	AAV67032 PERV env protein c
15	91	3.1	7362	22	AAV67019 PERV env protein c
16	91	3.1	7873	22	AAV67023 PERV env protein c
17	91	3.1	8060	18	AAV74811 Porcine retrovirus
18	91	3.1	8060	22	AAV77725 Tsukuba-1 cDNA. P
19	80	2.7	7892	18	AAV74884 Miniature swine re
20	74	2.5	4402	22	AAV67020 PERV env protein c
21	31	1.0	2702	19	AAV55872 Plant acetolactate
22	30	1.0	1431	21	AAV77334 Human ORFX ORF2889
23	30	1.0	1880	20	PERV-1-15 env prot
24	29	1.0	3557	7	Human pro-growth h
25	28	0.9	756	22	AAV521329 Human cDNA sequenc
26	28	0.9	925	10	AAV93196 Monkey recombinant
27	28	0.9	1493	21	AAV51086 3' DNA of porcine e
28	28	0.9	1736	21	AAV61203 cDNA encoding a hu
29	28	0.9	2000	21	AAV51087 Contiguous DNA of
30	28	0.9	10682	21	AAV66550 Human kinasin-like
31	27	0.9	327	21	AAV31562 Plant microsattel1
32	27	0.9	540	22	AAV26253 Pig endogenous ret
33	27	0.9	579	21	AAV50898 Wheat PI protein m
34	27	0.9	599	22	AAV26255 Pig endogenous ret
35	27	0.9	599	22	AAV26256 Pig endogenous ret
36	27	0.9	599	22	AAV26258 Pig endogenous ret
37	27	0.9	599	22	AAV26259 Pig endogenous ret
38	27	0.9	599	22	AAV26260 Pig endogenous ret
39	27	0.9	600	22	AAV26251 Pig endogenous ret
40	27	0.9	600	22	AAV26252 Pig endogenous ret
41	27	0.9	600	22	AAV26254 Pig endogenous ret
42	27	0.9	600	22	AAV26257 Pig endogenous ret
43	27	0.9	682	16	AAV02620 Met-des(F1-P2-F3-I
44	27	0.9	682	16	AAV11228 Met-des(F1-P2-F3-I
45	27	0.9	700	16	AAV02619 Met-des(Phel-Pro2-

ALIGNMENTS

RESULT 1	AAV09700	standard; DNA; 8209 BP.
ID	AAV09700	
AC	AAV09700;	
XX	20-MAR-1998	(first entry)
DT		
XX		
DE		Porcine retrovirus DNA encoding the LTR and GAG, POL and ENV proteins.
KW		Porcine retrovirus; POFV; POL protein; ENV protein; GAG protein;
KW		virion core polypeptide; polymerase protein; envelope protein;
KW		vaccine; diagnosis; xenotransplantation; prophylactic; therapeutic; ds.
OS		Porcine retrovirus.
XX		
XX		
Key	Location/Qualifiers	
FT	misc-feature	1..61
FT		/tag= a
FT		/note= "R LTR domain"
FT	LTR	1..588
FT		/tag= b
FT	misc-feature	62..143
FT		/tag= c
FT		/note= "U5 LTR domain"
FT	CDS	588..2162
FT		/tag= d
FT		/product= GAG protein
FT		/note= "viral core polypeptide"
FT	CDS	2163..5747
FT		/tag= e
FT		/product= POL protein
FT		/note= "polymerase peptide"
FT	CDS	5620..7590

```

FT      /tag= f
FT      /product= ENV protein
FT      /note= "envelope protein"
FT      7591..8209
FT      LTR
FT      /tag= g
FT      7638..8106
FT      /tag= h
FT      /note= "u3 LTR domain"
FT      7643..7686
FT      /tag= i
FT      /note= "Binding site for transcription factors
FT      Ocr-1, c-Myb, Lyf-1, E47 and ETS-1"
FT      7719..7739
FT      /tag= j
FT      /note= "Binding site for transcription factors
FT      Ap-4 and NF1"
FT      7744..7754
FT      /tag= k
FT      /note= "Binding site for transcription factor AP-1/TR"
FT      7773..7785
FT      /tag= l
FT      /note= "Binding site for transcription factor ETS-1/GATA"
FT      7793..7834
FT      /tag= m
FT      /note= "Binding site for transcription factors
FT      GATA, ETS-1, c-Myb, AP-1, GATA, E47"
FT      7841..7851
FT      /tag= n
FT      /note= "Binding site for transcription factor AP-1"
FT      7873..7883
FT      /tag= o
FT      /note= "Binding site for transcription factor AP-1"
FT      7885..7914
FT      /tag= p
FT      /note= "Binding site for transcription factors
FT      AP-1/GATA, c-Myb and AP-1"
FT      7928..7953
FT      /tag= q
FT      /note= "Binding site for transcription factors E47
FT      and ETS-1"
FT      7964..7968
FT      /tag= r
FT      7985..7997
FT      /tag= s
FT      /note= "Binding site for transcription factor GATA"
FT      8003..8007
FT      /tag= t
FT      8024..8036
FT      /tag= u
FT      /note= "Binding site for transcription factor GATA"
FT      8042..8046
FT      /tag= v
FT      8049..8060
FT      /tag= w
FT      /note= "Binding site for transcription factor
FT      AP-1/CRMB"
FT      8075..8090
FT      /tag= x
FT      1..61
FT      /tag= y
FT      /note= "R LTR domain"
FT      8127..8135
FT      /tag= z
FT      /note= "PolyA downstream element"
FT      8161..8166
FT      /tag= aa
XX      WO9740167-A1.
XX      30-OCF-1997.
XX      18-APR-1997;
XX      97WO-GB01087.

```

```

PR      10-FEB-1997; 97GB-0002668.
PR      19-APR-1996; 96GB-0008164.
XX      (IMUT-) IMUTRAN LTD.
PA      (QONE-) Q-ONE BIOTECH LTD.
XX      Galbraith DN, Haworth C, Lees GM, Smith KT;
XX      WPI: 1997-535851/49.
XX      P-PSDB: AAM39271; AAM39272; AAM39273.
XX      P-PSDB: AAM39271; AAM39272; AAM39273.
XX      This DNA sequence encodes the porcine retrovirus (PoEV) virion core
XX      polypeptide (GAG), polymerase (POL) and envelope (ENV) proteins and
XX      also includes the long terminal repeat (LTR). These proteins can be used
XX      to develop viral vaccines, antisense nucleic acids, ribozymes and other
XX      antiviral agents. They can also be used in xeno-transplantation
XX      technology and as diagnostic tools.
XX      Claim 4; Fig 3; 69pp; English.
XX      Sequence 8209 BP; 2168 A; 2064 C; 2154 G; 1823 T; 0 other:
SQ
Query Match          74.1%; Score 2191; DB 18; Length 8209;
Best local similarity 99.7%; Pred. No. 0;
Matches 2741; Conservative 0; Mismatches 5; Indels 3; Gaps 2;
QY      208 gagactgcaagttccacatcgctccaagttggagatcagttatgtttagagccacc 267
DB      5464 gagactgcaagttccacatcgctccaagttggagatcagttatgtttagagccacc 5523
QY      268 gtgcaggaaactcgaagctcggtggagaggaaccttacttgaacctgacacacaa 327
DB      5524 gtgcaggaaactcgaagctcggtggagaggaaccttacttgaacctgacacacaa 5583
QY      328 cggctgtgaaagtcgaagaaatctccaccttgatcatatgcatccacgttaagctggcgc 387
DB      5584 cggctgtgaaagtcgaagaaatctccaccttgatcatatgcatccacgttaagctggcgc 5643
QY      388 cactcccgactcgggtgtgagagcgcgaagaagactggaatcccttaagtttcctcc 447
DB      5644 cactcccgactcgggtgtgagagcgcgaagaagactggaatcccttaagtttcctcc 5703
QY      448 atcgctgttcttactactaaacaaataactcccgagcagtagtaaacgcttatagac 507
DB      5704 atcgctgttcttactactaaacaaataactcccgagcagtagtaaacgcttatagac 5763
QY      508 agctcgaaaccccatagacacttaccctacactgctgattatgacctgatacgggt 567
DB      5764 agctcgaaaccccatagacacttaccctacactgctgattatgacctgatacgggt 5823
QY      568 gtcactgttaataagaccacgaggtgtgtccttaagagcacctggtgcccgaactgcat 627
DB      5824 gtcactgttaataagaccacgaggtgtgtccttaagagcacctggtgcccgaactgcat 5883
QY      628 tctcgctcgaattatlaaccccgctgttaaaagcacacactcccaactagtcctagtc 687
DB      5884 tctcgctcgaattatlaaccccgctgtgttaaaagcacacactcccaactagtcctagtc 5943
QY      688 tatgggttctatgtcgtccagagcagagaagaataactgtgggttcttggggtaa 747
DB      5944 tatgggttctatgtcgtccagagcagagaagaataactgtgggttcttggggtaa 6003
QY      748 tctctctgtagagatgtagctggtcgaactcccaacagatgagagactggaatgagcgaatc 807
DB      6004 tctctctgtagagatgtagctggtcgaactcccaacagatgagagactggaatgagcgaatc 6063
QY      808 tctctcgaagaccggtlaaaattctccttctgtcaattccggtccggtgaagtacaaatg 867
DB      tctctcgaagaccggtlaaaattctccttctgtcaattccggtccggtgaagtacaaatg 867

```

Dh 6064 tctccaagaccggttaaaattctccttctgtaattccgccccgggcaagtaacaaatg 6123
Qy 868 atgaactataaagaataagactgtctcccaatccagacttagattatcttaagataagt 927
Dh 6124 atgaactataaagaataagactgtctcccaatccagacttagattatcttaagataagt 6183
Qy 928 ttcaatgaaaaaagaaaaacaggaataatctcaaaagtgsataaagtatgtagctggga 987
Dh 6184 ttccactg - aaggaataacaggaataatctcaaaagtgsataaagtatgtagctggga 6241
Qy 988 ataatttttaataataatgccccggggagcggttccactttaaccattcgcttaagata 1047
Dh 6242 ataattttttt-ataataaggcgggggagcggttccactttaaccattcgcttaagata 6300
Qy 1048 gagaacgggagacagacccccctgtgacatggaacccataaagtactggctgaacaggga 1107
Dh 6301 gagacggggagacagacccccctgtgacatggaacccataaagtactggctgaacaggga 6360
Qy 1108 cccccggccttgagccacacgcatlaacttgccggtgccccaaatlaacctgcgtcgccct 1167
Dh 6361 cccccggccttgagccacacgcatlaacttgccggtgccccaaatlaacctgcgtcgccct 6420
Qy 1168 gacatacacagccgcttaacaacagtaacacactggtatgattctcctacaacagcctaga 1227
Dh 6421 gacatacacagccgcttaacaacagtaacacactggtatgattctcctacaacagcctaga 6480
Qy 1228 aactccccaggtgtctcctgttaagacaggaacagactcttcagcttcacccaggagct 1287
Dh 6481 aactccccaggtgtctcctgttaagacaggaacagactcttcagcttcacccaggagct 6540
Qy 1288 ttccaagccatacactccacgacccctgatactgccaactctctgtgtgcttggctatcc 1347
Dh 6541 ttccaagccatacactccacgacccctgatactgccaactctctgtgtgcttggctatcc 6600
Qy 1348 tcagggcctcctattatagtgggggtatggtcctaagaagaataatctcaatgtgccaagag 1407
Dh 6601 tcagggcctcctattatagtgggggtatggtcctaagaagaataatctcaatgtgccaagag 6660
Qy 1408 catagaataatcatgtaacatggtggtcccgaaataagcttaacccctcaactgaagtctccgg 1467
Dh 6661 catagaataatcatgtaacatggtggtcccgaaataagcttaacccctcaactgaagtctccgg 6720
Qy 1468 aagggagacatgcataggaanaagctcccccatacccaacaacactgtgcatatgactgtg 1527
Dh 6721 aagggagacatgcataggaanaagctcccccatacccaacaacactgtgcatatgactgtg 6780
Qy 1528 gttatagagcagcctccagaataatcgtattatgtaacccgtgtatataacaggtgtggca 1587
Dh 6781 gttatagagcagcctccagaataatcgtattatgtaacccgtgtgtatataacaggtgtggca 6840
Qy 1588 tgcataactggttaacccccctgtgttccaactcagcttccaacaaatccaaagtgtg 1647
Dh 6841 tgcataactggttaacccccctgtgttccaactcagcttccaacaaatccaaagtgtg 6900
Qy 1648 tgtgtcatggttccaaatcgttcccccgagtgtaactaccatccagggaagtgtgtccctgat 1707
Dh 6901 tgtgtcatggttccaaatcgttcccccgagtgtaactaccatccagggaagtgtgtccctgat 6960
Qy 1708 gaataagcatatcgttataacggaacaaagaagaaacccgtatcccttaacccactgta 1767
Dh 6961 gaataagcatatcgttataacggaacaaagaagaaacccgtatcccttaacccactgta 7020
Qy 1768 atgctcgattaggaacgccccgtgtgctgafagaagaacaggaacggtgccccgtatcaaga 1827
Dh 7021 atgctcgattaggaacgccccgtgtgctgafagaagaacaggaacggtgccccgtatcaaga 7080
Qy 1828 ccacagcagctagagaagaagacttgggtgactatacatgctggccatgacagaagaatctccga 1887
Dh 7081 ccacagcagctagagaagaagacttgggtgactatacatgctggccatgacagaatctccga 7140
Qy 1888 gacctaaaggagctgtgttgcaacctagaagagctccctgaactcttctgtctgaagtgtt 1947
Dh 7141 gacctaaaggagctgtgttgcaacctagaagagctccctgaactcttctgtctgaagtgtt 7200

Qy 1948 ctacaagaccggagggatagatctgctgttctcaagaagagtggttatgtgcaacc 2007
Dh 7201 ctacagaaacccggagggatagatctgctgttctcaagaagagtggttatgtgcaacc 7260
Qy 2008 ttanaagaagaatgtgtctctctatgtatgatacactcaagagacatacagaactcatac 2067
Dh 7261 ttanaagaagaatgtgtctctctatgtatgatacactcaagagacatacagaactcatac 7320
Qy 2068 aagcttagaanaaagttaagagagcggtcgaaaggaaggaagcgaccagggtgtgtt 2127
Dh 7321 aagcttagaanaaagttaagagagcggtcgaaaggaaggaagcgaccagggtgtgtt 7380
Qy 2128 gaaggaatgttcaacagagctctctctgtgataagacacccctctctgctctgacgggccc 2187
Dh 7381 gaaggaatgttcaacagagctctctctgtgataagacacccctctctgctctgacgggccc 7440
Qy 2188 ctagtatgctctgctctgttacttaacaglttgggctgtgcttaataatagttgtgccc 2247
Dh 7441 ctagtatgctctgctctgttacttaacaglttgggctgtgcttaataatagttgtgccc 7500
Qy 2248 ttgtttaggaacgggtgagtgagtgagtgccagatcatalggttactaggaacagatcacaaagc 2307
Dh 7501 ttgtttaggaacgggtgagtgagtgagtgccagatcatalggttactaggaacagatcacaaagc 7560
Qy 2308 cctctgagccaaagggaacatgacctctagccttccagcttccagttctaagattagactat 2367
Dh 7561 cctctgagccaaagggaacatgacctctagccttccagcttccagttctaagattagactat 7620
Qy 2368 caagacaagaatgtgggaatgaaagatgaaatgcaacatlaacccctcccaagaacccag 2427
Dh 7621 caagacaagaatgtgggaatgaaagatgaaatgcaacatlaacccctcccaagaacccag 7680
Qy 2428 aagttataataaagctctaaatgcccccggaattacaagacctgctgctgtccagtaata 2487
Dh 7681 aagttataataaagctctaaatgcccccggaattacaagacctgctgctgtccagtaata 7740
Qy 2488 ggtagaaggttcaacttctctatgtgtccagggccgtgctatccctggccaaatlaaataac 2547
Dh 7741 ggtagaaggttcaacttctctatgtgtccagggccgtgctatccctggccaaatlaaataac 7800
Qy 2548 aggaataatgattgactaactgcttactgtgacttctgtaaaactgactgacacatagag 2607
Dh 7801 aggaataatgattgactaactgcttactgtgacttctgtaaaactgactgacacatagag 7860
Qy 2608 aattgattacacattgacagccctagtgatcctatctcaactgcaactcgtcaactctgcc 2667
Dh 7861 aattgattacacattgacagccctagtgatcctatctcaactgcaactcgtcaactctgcc 7920
Qy 2668 aggaagcccaagagatgagcagctcccggaactttaaagtatgtgtccaggaagc 2727
Dh 7921 aggaagcccaagagatgagcagctcccggaactttaaagtatgtgtccaggaagc 7980
Qy 2728 gggcctcgtatatttaaaatgattggtgcatactgagcgccgggctctcgatatttaaat 2787
Dh 7981 gggcctcgtatatttaaaatgattggtgcatactgagcgccgggctctcgatatttaaat 8040
Qy 2788 gattgttltgtgacgacagagcttgtgtgtgaaccccaataaagctgtgccgattccgca 2847
Dh 8041 gattgttltgtgacgacagagcttgtgtgtgaaccccaataaagctgtgtccgattccgca 8100
Qy 2848 ctgagggccggaatgctcttaccctcggtggtgtgtgtagagcttgggccccacgagcttgg 2907
Dh 8101 ctgagggccggaatgctcttaccctcggtggtgtgtgtagagcttgggccccacgagcttgg 8160
Qy 2908 aataaaaaatccctctgctgttgcatacaaaaaaataaaaaaataa 2956
Dh 8161 aataaaaaatccctctgctgttgcatacaaaaaaataaaaaaataa 8209

RESULT 2
AAV09699
ID AAV09699 standard; DNA; 8196 BP.

XX AAV09699;
AC 19-MAY-1998 (first entry)
XX Porcine retrovirus DNA encoding, GAG, POL and ENV.
XX Porcine retrovirus: PoEV; POL protein; ENV protein; GAG protein;
XX vaccine; diagnosis; xenotransplantation; prophylactic; therapeutic; ds.
XX Porcine retrovirus.
XX Location/Qualifiers
FH 576..2126
FT CDS
FT /tag= a
FT /product= GAG protein
FT /note= "viral core polypeptide"
FT 213..5733
FT /tag= b
FT /product= POL protein
FT /note= "polymerase peptide sequence as given in
FT specification"
FT 5606..7576
FT CDS
FT /tag= c
FT /product= ENV protein
FT /note= "envelope protein"
XX MO9740167-A1.
XX 30-OCT-1997.
XX 18-APR-1997; 97WO-GB01087.
XX 10-FEB-1997; 97GB-0002668.
XX 19-APR-1996; 96GB-0008164.
XX (IMUT-) IMUTRAN LTD.
XX (OONE-) O-ONE BIOTECH LTD.
XX Galbraith DN, Haworth C, Lees GM, Smith KT;
XX WPI; 1997-535851/49.
XX Polynucleotide encoding porcine retrovirus expression product -
XX useful to develop products for use in vaccines, diagnosis and
XX xeno-transplantation
XX Claim 4; Fig 2; 69pp; English.
XX This DNA sequence encodes the porcine retrovirus (PoEV) virion core
XX polypeptide (GAG), polymerase (POL) and envelope (ENV) proteins. These
XX proteins can be used to develop viral vaccines, antisense nucleic acids,
XX ribozymes and other antiviral agents. They can also be used in
XX xeno-transplantation technology and as diagnostic tools.
XX Sequence 8196 BP; 2165 A; 2061 C; 2147 G; 1820 T; 3 other;
Query Match 70.7%; Score 2089; DB 18; Length 8196;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2739; Conservative 0; Mismatches 7; Indels 3; Gaps 2;

QY 388 caactccgactcgggtgagagcgaagaagactgaaatcccttaagcttcgctcc 447
DB 5630 caactccgactcgggtgagagcgaagaagactgaaatcccttaagcttcgctcc 5689
QY 448 atcgccgtgtcccttaacttaacaataactcccgagccagtagtaaacgctctaaac 507
DB 5690 atcgccgtgtcccttaacttaacaataactcccgagccagtagtaaacgctctaaac 5749
QY 508 agctggaaccaccatagactcttacccttaactgctgcatattgaccctgatacgggt 567
DB 5750 agctggaaccaccatagactcttacccttaactgctgcatattgaccctgatacgggt 5809
QY 568 gtcaactgtaaatagactcgaagtggtgtcccttaagagcaactgtgtgacctgaat 627
DB 5810 gtcaactgtaaatagactcgaagtggtgtcccttaagagcaactgtgtgacctgaat 5869
QY 628 ttcgctcccgattgattaaacccgctgttaaaacacacccccaactgctcgtatgt 687
DB 5870 ttcgctcccgattgattaaacccgctgttaaaacacacccccaactgctcgtatgt 5929
QY 688 tatgggttctatgtcgtccaggaagcaagaagaataactgtggtggttctgaggaa 747
DB 5930 tatgggttctatgtcgtccaggaagcaagaagaataactgtggtggttctgaggaa 5989
QY 748 tccctctgtagagatgagatgagctgctcaccctcaacatgagactggaatgagcagatc 807
DB 5990 tccctctgtagagatgagatgagctgctcaccctcaacatgagactggaatgagcagatc 6049
QY 808 tctctcgaagacccgggttaaatcttcccttgcattccgcccggcggaagtagaataatg 867
DB 6050 tctctcgaagacccgggttaaatcttcccttgcattccgcccggcggaagtagaataatg 6109
QY 868 atgaactatataaagataagactgctcccccatacagacttgaattatctaaagaatagt 927
DB 6110 atgaactatataaagataagactgctcccccatacagacttgaattatctaaagaatagt 6169
QY 928 ttcactgaaagaagaagaagaataatctcaaaagtgtataatgtatgtatgctcgggga 987
DB 6170 ttcactg--aaagaaagaagaagaataatctcaaaagtgtataatgtatgtatgctcgggga 6227
QY 988 atagttttttaaataatgacgggggagagaggttcaacttcaacttgccttgctgagata 1047
DB 6228 atagttttt-attataatgacgggggagagaggttcaacttcaacttgccttgctgagata 6286
QY 1048 gagacgggagacgaagaacccctgttgcaatgagacccgaataaagtactggtcgaacagg 1107
DB 6287 gagacgggagacgaagaacccctgttgcaatgagacccgaataaagtactggtcgaacagg 6346
QY 1108 ccccgccctgagacccacgataactgtccggtggtcccaattcaactcgtctgctgacct 1167
DB 6347 ccccgccctgagacccacgataactgtccggtggtcccaattcaactcgtctgctgacct 6406
QY 1168 gacataacacagccgctgagacacagatgacactgtgattatccctcaacacagcctaga 1227
DB 6407 gacataacacagccgctgagacacagatgacactgtgattatccctcaacacagcctaga 6466
QY 1228 aactcccaagtgctcgttlaagacagagacagactctcaagtcacatccaggaagct 1287
DB 6467 aactcccaagtgctcgttlaagacagagacagactctcaagtcacatccaggaagct 6526
QY 1288 ttccaagcacaactcaacccgacccctgtatgcaactctcttgtgtgtgtgtctatcc 1347
DB 6527 ttccaagcacaactcaacccgacccctgtatgcaactctcttgtgtgtgtgtctatcc 6586
QY 1348 tcaagggtcccttattatagagagatgagctgtaagaagaagaataatcaatgtgacaaag 1407
DB 6587 tcaagggtcccttattatagagagatgagctgtaagaagaagaataatcaatgtgacaaag 6646
QY 1408 catagaatcaatgatacaggggttcccgaaataaacttaacctcaactgaagtcccg 1467
DB 6647 catagaatcaatgatacaggggttcccgaaataaacttaacctcaactgaagtcccg 6706

CC isolated from the human cell line Raji. Such viral proteins can be used
 CC to develop viral vaccines, antisense nucleic acids, ribozymes and other
 CC antiviral agents. They can also be used in xeno-transplantation
 CC technology and as diagnostic tools.

XX Sequence 1974 BP: 534 A; 489 C; 483 G; 468 T; 0 other;

Query Match 63.3%; Score 1872; DB 18; Length 1974;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1972; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 364 atgacccaccgcttaagctggcgccaccccccgcgctgggggtgagagccgcaaaagctg 423
 Db 1 atgacccaccgcttaagctggcgccaccccccgcgctgggggtgagagccgcaaaagctg 60
 QY 424 agaatcccttaagcttcgcctcctacgtcgtgttccttactctaaataaactccccag 483
 Db 61 agaatcccttaagcttcgcctcctacgtcgtgttccttactctaaataaactccccag 120
 QY 484 gccagtagtaaacgcttatagacagctcgaaaccccatagacctttatccctacgtg 543
 Db 121 gccagtagtaaacgcttatagacagctcgaaaccccatagacctttatccctacgtg 180
 QY 544 ctgattatgacccctgtagcgggtgtcactgttaaatagcactcgaagtgtgtcctaga 603
 Db 181 ctgattatgacccctgtagcgggtgtcactgttaaatagcactcgaagtgtgtcctaga 240
 QY 604 gccacccgtgtgctgtgaacttgccttcgcctccgattgattaaacccgctgttaaac 663
 Db 241 gccacccgtgtgctgtgaacttgccttcgcctccgattgattaaacccgctgttaaac 300
 QY 664 acaccccccacccctagctcgtgagctatggtgttcctatgctgcgcccagcacagagaaag 723
 Db 301 acaccccccacccctagctcgtgagctatggtgttcctatgctgcgcccagcacagagaaag 360
 QY 724 aaatactgtgggggtctctgggaaatcctctctgtagagatgagctgtgcacccccaac 783
 Db 361 aaatactgtgggggtctctgggaaatcctctctgtagagatgagctgtgcacccccaac 420
 QY 784 gatgagagcttggaaatggtccgcatctctctccagacccgggttaaattctcttgttaac 843
 Db 421 gatgagagcttggaaatggtccgcatctctctccagacccgggttaaattctcttgttaac 480
 QY 844 tccggcccgagcaagatgaaatgataataataaagctgtcccccata 903
 Db 481 tccggcccgagcaagatgaaatgataataataaagctgtcccccata 540
 QY 904 gacttagattactaaagataagtttcaactgaaagaaagaaacagagaaatattcaaaag 963
 Db 541 gacttagattactaaagataagtttcaactgaaagaaagaaacagagaaatattcaaaag 600
 QY 964 tggataaatgtatgagctgtgggaaatgttttataataatgtgctggggagacaggttcc 1023
 Db 601 tggataaatgtatgagctgtgggaaatgttttataataatgtgctggggagacaggttcc 660
 QY 1024 acttaacacatcgctcttagatagagacgggacagaaacccctgtggaatgagagccc 1083
 Db 661 acttaacacatcgctcttagatagagacgggacagaaacccctgtggaatgagagccc 720
 QY 1084 gataaagtactgtgtgaacaggggcccgcgctgtgagccacgcgataactgtgcccgtg 1143
 Db 721 gataaagtactgtgtgaacaggggcccgcgctgtgagccacgcgataactgtgcccgtg 780
 QY 1144 ccccaatlaacctgtgtgagctgtgacataacacagccgcttagcaacagttaccatgtga 1203
 Db 781 ccccaatlaacctgtgtgagctgtgacataacacagccgcttagcaacagttaccatgtga 840
 QY 1204 ttgattcttaccacacgcttagaagaactcccaaggtgttctgtttaagaagaagagaga 1263
 Db 841 ttgattcttaccacacgcttagaagaactcccaaggtgttctgtttaagaagaagagaga 900
 QY 1264 ctcttcagctcatccagggagcttccaaagccatcaactcaacgcgacctgtatgacct 1323

Db 901 cttctacgtcatccagggagcttccaaagccatcaactcaacgcgacctgatgacct 960
 QY 1324 tctctgtgtgtgtgtgtatccatccagggcctctcttattatgagggaggtgttaaaag 1383
 Db 961 tctctgtgtgtgtgtgtatccatccagggcctctcttattatgagggaggtgttaaaag 1020
 QY 1384 agaaaaatcaatgtgacccaagaagacataaataatgataatgtgggggtccccaataag 1443
 Db 1021 agaaaaatcaatgtgacccaagaagacataaataatgataatgtgggggtccccaataag 1080
 QY 1444 cttaacctactgaagttccgggaaaggagacatgacataggaagaagctcccccatacc 1503
 Db 1081 cttaacctactgaagttccgggaaaggagacatgacataggaagaagctcccccatacc 1140
 QY 1504 caaaccttgcctatgactgtgtgttattgagcagggcctcagaagaatcaagratattga 1563
 Db 1141 caaaccttgcctatgactgtgtgttattgagcagggcctcagaagaatcaagratattga 1200
 QY 1564 cctgtattacaacaggtgtggtgacatgcaatactgggttlaacccctgtgttccacctca 1623
 Db 1201 cctgtattacaacaggtgtggtgacatgcaatactgggttlaacccctgtgttccacctca 1260
 QY 1624 gtcttcaaccaaaccacaagaattgtgtgtcatgtgtccaaatcgctcccgagtgataac 1683
 Db 1261 gtcttcaaccaaaccacaagaattgtgtgtcatgtgtccaaatcgctcccgagtgataac 1320
 QY 1684 catcctgaaagagtggtctcttattgataatgactctcgttatcaacccgcaaaagaa 1743
 Db 1321 catcctgaaagagtggtctcttattgataatgactctcgttatcaacccgcaaaagaa 1380
 QY 1744 cccgtatcccttaacccctacgttaatgtcctcagatctagagagccgctgtggtcgtlaagaa 1803
 Db 1381 cccgtatcccttaacccctacgttaatgtcctcagatctagagagccgctgtggtcgtlaagaa 1440
 QY 1804 gggacagctgcccctgtatcaacaggaaccaagcagctagagaagaagctgtgtgactaac 1863
 Db 1441 gggacagctgcccctgtatcaacaggaaccaagcagctagagaagaagctgtgtgactaac 1500
 QY 1864 gggcccatagacagaagaatcccgagccttaagaagagctcgttagaacaactaagaagctc 1923
 Db 1501 gggcccatagacagaagaatcccgagccttaagaagagctcgttagaacaactaagaagctc 1560
 QY 1924 ctgacttctgtctgaagtggtttctacagaacccggagggagtttagatctgtgttcta 1983
 Db 1561 ctgacttctgtctgaagtggtttctacagaacccggagggagtttagatctgtgttcta 1620
 QY 1984 agagaagtggtgtatgtgacgctttaaagaagaatgttgcctctatgtatcaactca 2043
 Db 1621 agagaagtggtgtatgtgacgctttaaagaagaatgttgcctctatgtatcaactca 1680
 QY 2044 ggaagccatagagactccctgaacagctttaaagaagaatgttgaagagcgttgaaggaa 2103
 Db 1681 ggaagccatagagactccctgaacagctttaaagaagaatgttgaagagcgttgaaggaa 1740
 QY 2104 agagaagctgacaaaggggtgtgttgaagatgttcaacaagctcctctgtatgacaccc 2163
 Db 1741 agagaagctgacaaaggggtgtgttgaagatgttcaacaagctcctctgtatgacaccc 1800
 QY 2164 ctgcttctgtctgtgacgggcccctagtagctcgtctcctgttacttaacagtgtggcct 2223
 Db 1801 ctgcttctgtctgtgacgggcccctagtagctcgtctcctgttacttaacagtgtggcct 1860
 QY 2224 tgccttaataataggtttgttgccttctgttgaagaagagatgagtgacgttcaagatatg 2283
 Db 1861 tgccttaataataggtttgttgccttctgttgaagaagagatgagtgacgttcaagatatg 1920
 QY 2284 gtacttagcaacagatcaacaagccttctgaacaaagagaaactgacctctag 2337
 Db 1921 gtacttagcaacagatcaacaagccttctgaacaaagagaaactgacctctag 1974

RESULT 4

AAV82749
 ID AAV82749 standard; DNA; 3482 BP.
 AC AAV82749;
 XX
 XX
 XX 25-FEB-1999 (first entry)
 XX
 XX Pig endogenous retrovirus (PERV)-B envelope (env) gene region.
 DE
 XX Pig endogenous retrovirus; PERV-A; envelope protein; PERV-B; subtype;
 KW probe; detection; retrovirus; human tissue; xenotransplant;
 KW primary porcine tissue; human cell line; porcine cell line; ss.
 XX
 XX Pig endogenous retrovirus.
 OS
 XX
 XX
 XX Key Location/Qualifiers
 FT 911.2884
 FT /*tag= a
 FT /product= envelope_protein
 XX
 XX WO9853104-A2.
 XX
 XX 26-NOV-1998.
 XX
 XX 18-MAY-1998; 98WO-GB01428.
 XX
 XX 16-MAY-1997; 97GB-0010154.
 XX
 XX (MED1-) MEDICAL RES COUNCIL.
 XX
 XX Stoye JP, Weiss RA;
 PI
 XX
 XX MPI; 1999-045324/04.
 DR P-PSDB; AAW85452.
 XX
 XX Newly isolated nucleic acid probe capable of hybridising to either
 PT the PERV-A or PERV-B env gene - useful in the detection of
 PT retroviruses, and their subtypes, in a sample of porcine/human
 PT tissue
 XX
 XX
 XX Claim 3: Page 21-23; 36pp: English.
 PS
 XX The present sequence encodes a pig endogenous retrovirus (PERV)-B
 CC envelope protein. PERV exists in two different subtypes, PERV-A and
 CC PERV-B. The differences are reflected in sequence divergence in the
 CC envelope genes. Probes and primers can be derived from the envelope
 CC (env) genes of PERV-A and PERV-B. The probes and primers are used in
 CC a method to detect retroviruses in a sample of porcine/human tissue,
 CC particularly primary porcine tissue and human cell lines that have been
 CC cultivated in the presence of a porcine cell line, or human tissue from
 CC a patient with a xenotransplant. Subtype of PERV in a sample containing
 CC one of the PERV env genes can also be determined.
 CC
 CC
 CC Sequence 3482 BP; 927 A; 854 C; 867 G; 834 T; 0 other:
 SQ

Query Match 59.0%; Score 1743; DB 20; Length 3482;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 2913; Conservative 0; Mismatches 21; Indels 1; Gaps 1;

Db 728 agctccggagagcctactcagagagagagacttgcaagttccactcgtcccaagtgg 787
 QY 241 gagatcagctatgttagacgcaccgltgcaagaaacctcgagactcgttggaagggac 300
 Db 788 gagatcagctatgttagacgcaccgltgcaagaaacctcgagactcgttggaagggac 847
 QY 301 ctatctcgtacttttgaccacacaaagcgtgtgtaagtgcgaagaaatccacttga 360
 Db 848 ctatctcgtacttttgaccacacaaagcgtgtgtaagtgcgaagaaatccacttga 907
 QY 361 tccatgcatccacglttaagcgtgcgcacacctccgacccggggtgagagccgaaaga 420
 Db 908 tccatgcatccacglttaagcgtgcgcacacctccgacccggggtgagagccgaaaga 967
 QY 421 ctgagaaaccccttaagcttcgctccatcgcctgtgttcttacttcaataactccc 480
 Db 968 ctgagaaaccccttaagcttcgctccatcgcctgtgttcttacttcaataactccc 1027
 QY 481 caggccagtagtaaacgaccttaagcagctggaaccccatagaccttaccctaac 540
 Db 1028 caggccagtagtaaacgaccttaagcagctggaaccccatagaccttaccctaac 1087
 QY 541 tggctgattatgacccctgatacgggtgctcactgttaaatagcagcgaagtgtgctct 600
 Db 1088 tggctgattatgacccctgatacgggtgctcactgttaaatagcagcgaagtgtgctct 1147
 QY 601 agagcagcctgtggtgcctgaactgcattctgcctccgattgataaacccgctgttaa 660
 Db 1148 agagcagcctgtggtgcctgaactgcattctgcctccgattgataaacccgctgttaa 1207
 QY 661 agcacacctcccaacctagctccgttagttgtgtctatgtctgcgccggacagagaa 720
 Db 1208 agcacacctcccaacctagctccgttagttgtgtctatgtctgcgccggacagagaa 1267
 QY 721 gagaatactgtggtggtcttggtggaatcctctctgtagagatgagctgcgtcaccctc 780
 Db 1268 gagaatactgtggtggtcttggtggaatcctctctgtagagatgagctgcgtcaccctc 1327
 QY 781 aacgatgagagactggaatlgccgatctctccagagaccgggtaaattctccttctc 840
 Db 1328 aacgatgagagactggaatlgccgatctctccagagaccgggtaaattctccttctc 1387
 QY 841 aattccggccggggaagaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaa 900
 Db 1388 aattccggccggggaagaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaa 1447
 QY 901 tcaagcttagattcctaagaataglttcaactgtaaaagaaagaaagaaatattcaa 960
 Db 1448 tcaagcttagattcctaagaataglttcaactgtaaaagaaagaaagaaatattcaa 1507
 QY 961 aagtgtgataaagtgtatgagctggtggaatagtttttaataatcgtgcggggagcagg 1020
 Db 1508 aagtgtgataaagtgtatgagctggtggaatagtttttaataatcgtgcggggagcagg 1567
 QY 1021 tccactttaaaccttcgcttagagtagagagagagagagagagagagagagagagagag 1080
 Db 1568 tccactttaaaccttcgcttagagtagagagagagagagagagagagagagagagagag 1627
 QY 1081 ccggaataaagtaactgctgtaaaagggcccgccggtgacacgcgcataacttgcg 1140
 Db 1628 ccggaataaagtaactgctgtaaaagggcccgccggtgacacgcgcataacttgcg 1687
 QY 1141 gtagcccaatfaactcgtcgtgcgcctgacataacacagccgctagcaacagtaacact 1200
 Db 1688 gtagcccaatfaactcgtcgtgcgcctgacataacacagccgctagcaacagtaacact 1747
 QY 1201 gtagtattcctacacacagccctagaaactcccaagggttctgtttaaagagagagag 1260
 Db 1748 gtagtattcctacacacagccctagaaactcccaagggttctgtttaaagagagagag 1807
 QY 1261 agactcttaagttcctacacagagagcttccaagccatcaactccacgagaccgtatgccc 1320

Db	1808	agactctcagctctcatccaaaggaggttccaaagcatatcaactccacccgagacctgagctc	1867
Qy	1321	actctctctgtgtggtctgtgtctatccctcaaggcctccttattatgagggagatggtctaa	1380
Db	1868	actctctctgtgtgtgtgtgtctatccctcaaggcctccttattatgagggagatggtctaa	1927
Qy	1381	gaaagaaattctaatgtgaccaaaggagctaatcaattcaattgactgtgggtctccgaat	1440
Db	1928	gaaggaagaattcaatgtgaccaaaggagctaatcaattcaattgactgtgggtctccgaat	1987
Qy	1441	aagtttaacctctcactgaagtttcccggaaggagacatgcatagaaaagctccccaatc	1500
Db	1988	aagtttaacctctcactgaagtttcccggaaggagacatgcatagaaaagctccccaatc	2047
Qy	1501	caccacaacctgtgtcatatgactgtgtgttttttttttttttttttttttttttttttttt	1560
Db	2048	caccacaacctgtgtcatatgactgtgtgttttttttttttttttttttttttttttttttt	2107
Qy	1561	gtacctgttatacaacagtgtygtgagctgacaatgcaactggttttaacccctgtgttccac	1620
Db	2108	gtacctgtgttatacaacagtgtygtgagctgacaatgcaactggttttaacccctgtgttccac	2167
Qy	1621	tcagttcttcaaccaatcccaagaattgtgtgtctgcatggttccaatcgctcccgagtgta	1680
Db	2168	tcagttcttcaaccaatcccaagaattgtgtgtctgcatggttccaatcgctcccgagtgta	2227
Qy	1681	taccatccctaggaagagtggtcctctgtatgaaatgactatcggtataacccgacaaaaga	1740
Db	2228	taccatccctaggaagagtggtcctctgtatgaaatgactatcggtataacccgacaaaaga	2287
Qy	1741	gaaccgctatccctcttaccctagctgtgaatgtctcgattagggagcggcttggtcgtaga	1800
Db	2288	gaaccgctatccctcttaccctagctgtgaatgtctcgattagggagcggcttggtcgtaga	2347
Qy	1801	acaaaggacagtggtcctctgactacacagagccaacagctagaagaagaagccttggttagta	1860
Db	2348	acaaaggacagtggtcctctgactacacagagccaacagctagaagaagaagccttggttagta	2407
Qy	1861	catctgcgcacatgacaagaagaattccctcgagcctttaaagagttctgttagaacacttagaag	1920
Db	2408	catctgcgcacatgacaagaagaattccctcgagcctttaaagagttctgttagaacacttagaag	2467
Qy	1921	tccctcgactcttctgtgttgtaagtggtttttaaagcaacccggaagggttttagatctgtgtt	1980
Db	2468	tccctcgactcttctgtgttgtaagtggtttttaaagcaacccggaagggttttagatctgtgtt	2527
Qy	1981	ctaaagaaaggttgagttgtatgtgacgcctttaaaagaagaattgtctgtctatgttagatcac	2040
Db	2528	ctaaagaaaggttgagttgtatgtgacgcctttaaaagaagaattgtctgtctatgttagatcac	2587
Qy	2041	tcaaggagcactcagagaagcttccatgacaagaagcttaagaaaaagttagagagagtgctcgaaag	2100
Db	2588	tcaaggagcactcagagaagcttccatgacaagaagcttaagaaaaagttagagagagtgctcgaaag	2647
Qy	2101	gaaagaaagagtgtagccaggggtgtgtttgaaagatagtttacaagagcttccctgtgatacc	2160
Db	2648	gaaagaaagagtgtagccaggggtgtgtttgaaagatagtttacaagagcttccctgtgatacc	2707
Qy	2161	aacctggtcttctgtctgtacaggggcccctgagtacgtccgtctccgttacttaacagtgtg	2220
Db	2708	aacctggtcttctgtctgtacaggggcccctgagtacgtccgtctccgttacttaacagtgtg	2767
Qy	2221	cctgtcttaattataatgagttgtgtgtcctgtgttaagaaacaggtgagtgcaatccagatc	2280
Db	2768	cctgtcttaattataatgagttgtgtgtcctgtgttaagaaacaggtgagtgcaatccagatc	2827
Qy	2281	atgtagtacttaagcaacagtaacaaagagccttctgagccaaaggagaaactctacagct	2340
Db	2828	atgtagtacttaagcaacagtaacaaagagccttctgagccaaaggagaaactctacagct	2887
Qy	2341	tcccaagttctaagatttagactattatacaagacaagaagtggtggataatgaagaatgaaaa	2400
Db	2888	tcccaagttctaagatttagactattatacaagacaagaagtggtggataatgaagaatgaaaa	2947

[illegible]

RESULT

ID AAF77726 standard; DNA; 7333 BP.

AC AAF77726;

DT 23-MAY-2001 (first entry)

DE Defective retroviral genome isolated from PK-15 cell line.

AA Retrovirus; graft transplantation; xenotransplantation; PK-15 cell line;
KW

X X
X X

5
5
.

unidentified
05
XX

056190861-BL
PN
XX

PD 20-FEB-2001.
YY

13-DEC-1996; 9605-0/66528.
PE
XX

PR 14-DEC-1995; 95US-0572645.

PA (GEHO) GEN HOSPITAL CORP.

PI Fishman JA;

DR WPI; 2001-256211/26.

XX

PT xenotransplantation, comprises using probe sequences derived from swine

XX

2

XX The present invention relates to a method for screening a cell or tissue
 CC for the presence or expression of a retrovirus (RV), comprising
 CC contacting a target nucleic acid from the cell or tissue with a second
 CC nucleic acid from the present invention (e.g. the present sequence or a
 CC fragment thereof). The method is useful for RV detection and to assess
 CC graft transplantation risk. Screening of animals allows the elimination
 CC of donors with active replication of known viruses. Inactive proviruses
 CC can be detected and inactivated, allowing identification and elimination
 CC of potential human pathogens derived from swine in a manner not possible
 CC in the outbred human organ donor population and is important to the
 CC development of human xenotransplantation.

SQ Sequence 7333 BP; 1984 A; 1788 C; 1885 G; 1676 T; 0 other;

Query Match 27.5%; Score 813; DB 22; Length 7333;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 863; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2082 gttagagagcgctcgaaggaagagagcgctaccaggggtgtttgaagatgttcaa 2141
 Db 6470 gttagagagcgctcgaaggaaggaagagcgctaccaggggtgtttgaagatgttcaa 6529
 OY 2142 caggtctccttgatgataccaccctgtcttctgtctctgaagggccctagatctctgt 2201
 Db 6530 caggtctccttgatgataccaccctgtcttctgtctctgaagggccctagatctctgt 6589
 OY 2202 cctgttacttaacagtgtggccttcttaataaggttctgtccttgttagaagacg 2261
 Db 6590 cctgttacttaacagtgtggccttcttaataaggttctgtccttgttagaagacg 6649
 OY 2262 agtgaatgagatccagatatagttacttaggaacagatccagagccttcttagcagaag 2321
 Db 6650 agtgaatgagatccagatatagttacttaggaacagatccagagccttcttagcagaag 6709
 OY 2322 agaaactgacacctagccttccagttcttaagattagaaacttaacagaacaaagtg 2381
 Db 6710 agaaactgacacctagccttccagttcttaagattagaaacttaacagaacaaagtg 6769
 OY 2382 ggggaatgaaagatgaaatgcaacctaaacctccagagaccaggaagttaataaag 2441
 Db 6770 ggggaatgaaagatgaaatgcaacctaaacctccagagaccaggaagttaataaag 6829
 OY 2442 cctaataatgccccgaataataacagacccgtgtgtcgcgaagtaaatagatgaagttcaca 2501
 Db 6830 cctaataatgccccgaataataacagacccgtgtgtcgcgaagtaaatagatgaagttcaca 6889
 OY 2502 cttcctatgttccagggcctgtctacctgtgcctaaagataacacagaaatgagttga 2561
 Db 6890 cttcctatgttccagggcctgtctacctgtgcctaaagataacacagaaatgagttga 6949
 OY 2562 ctaatgcgttatcctgattctgttaaaactgactgacacataagaagaattgattcacat 2621
 Db 6950 ctaatgcgttatcctgattctgttaaaactgactgacacataagaagaattgattcacat 7009
 OY 2622 tgaacagccctaaagtccatactcaactgaatcttgaactctgcccaggaagccccagag 2681
 Db 7010 tgaacagccctaaagtccatactcaactgaatcttgaactctgcccaggaagccccagag 7069
 OY 2682 atgcgagacctccgagactatttaaaatgattgtgccacgagcgcggtctcgatatc 2741
 Db 7070 atgcgagacctccgagactatttaaaatgattgtgccacgagcgcggtctcgatatc 7129
 OY 2742 ttaaaatgattgtccatagagcgcggtctcgatatatttaaaatgattgtgtgac 2801
 Db 7130 ttaaaatgattgtccatagagcgcggtctcgatatatttaaaatgattgtgtgac 7189
 OY 2802 gcaacagccttgttgaacccccataaaagctgtccgatttcgcagcggcgccgcagct 2861
 Db 7190 gcaacagccttgttgaacccccataaaagctgtccgatttcgcagcggcgccgcagct 7249
 OY 2862 ccttaacccctgtgtgtgtacgactgtgggcccagcgcgcttgaataaaatccctct 2921

Db 7250 ccttaacccctgtgtgtgtacgactgtgggcccagcgcgcttgaataaaatccctct 7309
 OY 2922 tgcgttgcatacaaaaaaaaaa 2945
 Db 7310 tgcgttgcatacaaaaaaaaaa 7333

RESULT 6

ID AAV82748 standard; DNA; 2462 BP.

AAV82748:

25-FEB-1999 (first entry)

DE Pig endogenous retrovirus (PERV)-A envelope (env) gene region.

KW Pig endogenous retrovirus; PERV-A; envelope protein; PERV-B; subtype;

KW probe; primer: detection; retrovirus; human tissue; xenotransplant;

XX primary porcine tissue; human cell line; porcine cell line; ss.

OS Pig endogenous retrovirus.

XX Key Location/Qualifiers

XX CDS 211..2193

XX FT //tag= a

XX FT /product= envelope-protein

XX MO9853104-A2.

XX PD 26-NOV-1998.

XX PE 18-MAY-1998; 98MO-GB01428.

XX PR 16-MAY-1997; 97GB-0010154.

XX PA (MED1-) MEDICAL RES COUNCIL.

XX PI Stoye JP, Weiss RA;

XX DR WPL: 1999-045324/04.

XX PS P-PSDB: AAM85453.

XX PT Newly isolated nucleic acid probe capable of hybridizing to either

XX PT the PERV-A or PERV-B env gene - useful in the detection of

XX PT retroviruses, and their subtypes, in a sample of porcine/human

XX PT tissue

XX PS Claim 6; Page 20-21; 36pp; English.

XX CC The present sequence encodes a Pig endogenous retrovirus (PERV)-A

XX CC envelope protein. PERV exists in two different subtypes, PERV-A and

XX CC PERV-B. The differences are reflected in sequence divergence in the

XX CC envelope genes. Probes and primers can be derived from the envelope

XX CC (env) genes of PERV-A and PERV-B. The probes and primers are used in

XX CC a method to detect retroviruses in a sample of porcine/human tissue,

XX CC particularly primary porcine tissue and human cell lines that have been

XX CC cultivated in the presence of a porcine cell line, or human tissue from

XX CC a patient with a xenotransplant. Subtype of PERV in a sample containing

one of the PERV env genes can also be determined.

XX SQ Sequence 2462 BP; 715 A; 573 C; 575 G; 599 T; 0 other;

Query Match 15.5%; Score 457; DB 20; Length 2462;

Best Local Similarity 99.8%; Pred. No. 2,6e-178;

Matches 507; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2082 gttagagagcgctcgaaggaagagagcgctaccaggggtgtttgaagatgttcaa 2141
 Db 1938 gttagagagcgctcgaaggaagagagcgctaccaggggtgtttgaagatgttcaa 1997

```

QY 2142 caagctctcttgatgacccaccctgcttctctctgacggggccctagatcctgct 2201
XX |||||||
XX 1998 caagctctcttgatgacccaccctgcttctctctgacggggccctagatcctgct 2057
QY 2202 cctgttactaacgttgggcttgccttaataatagttgttgccttctgttagaagc 2261
XX |||||||
XX 2058 cctgttactaacgttgggcttgccttaataatagttgttgccttctgttagaagc 2117
QY 2262 agtgaagcagctcagatcgttacttagcacaacgtacaaagggcttctgacaaag 2321
XX |||||||
XX 2118 agtgaagcagctcagatcgttacttagcacaacgtacaaagggcttctgacaaag 2177
QY 2322 agaaactgacctagccttccagttctaagattagaactataaacaagaagaagtg 2381
XX |||||||
XX 2178 agaaactgacctagccttccagttctaagattagaactataaacaagaagaagtg 2237
QY 2382 gggatgaagaagatgaataagcaacctaacctcccaagaagccaggaagtataaaga 2441
XX |||||||
XX 2238 gggatgaagaagatgaataagcaacctaacctcccaagaagccaggaagtataaaga 2297
QY 2442 ctctaaatgccccgaattacagacctgtgctgctgccaataatagatgaagtcaca 2501
XX |||||||
XX 2298 ctctaaatgccccgaattacagacctgtgctgctgccaataatagatgaagtcaca 2257
QY 2502 ctctctattgttccagggctctatcctgctgccaataatagatgaagaatgagttga 2561
XX |||||||
XX 2358 ctctctattgttccagggctctatcctgctgccaataatagatgaagaatgagttga 2417
QY 2562 ctatcgcttatctgagttctgttaaac 2589
XX |||||||
XX 2418 ctatcgcttatctgagttctgttaaac 2445

```

RESULT 7
AAT74883 standard; cDNA; 7393 BP.

AC AAT74883:
XX 09-FEB-1998 (first entry)
XX
XX Porcine retrovirus cDNA (defective).
DE
XX
XX Retrovirus; porcine; GAG protein; POL protein; ENV protein;
KW xenotransplantation; infectious; provirus; organ transplant; donor;
KW activated virus; PCR; ss.
XX
XX Porcine retrovirus.
OS
XX

Location/Qualifiers

FH Key 598..2172
FT CDS /tag= a
FT 598..2169
FT mat_peptide /tag= b
FT /note= "putative GAG protein"
FT 2320..4737
FT /tag= c
FT /note= "putative POL coding region (partial) as
FT 2320..3522 described in the specification"
FT /tag= d
FT /note= "putative POL protein (partial)"
FT 3516..4328
FT mat_peptide /tag= e
FT /note= "putative POL protein (partial)"
FT 4332..4748
FT mat_peptide /tag= f
FT 4738..6725
FT CDS /tag= g
FT /note= "putative ENV coding region (partial) as
FT 4752..6722 described in the specification"
FT /tag= h

```

FT /note= "ENV protein (partial)"
XX WO9721836-A1.
XX 19-JUN-1997.
XX 13-DEC-1996; 96WO-US19680.
XX 14-DEC-1995; 95US-0572645.
XX (GENO ) GEN HOSPITAL CORP.
XX
XX Fisiman JA;
XX WPI; 1997-332804/30.
XX P-PSDB; AAM32091-W32095.
XX
XX New nucleic acid from porcine retroviruses - used for detecting
XX PT viruses in transplant or other tissue and for assessing risk of
XX PT transmitting infection to graft recipient
XX
XX Claim 16; Fig 2; 128pp; English.
XX
XX This cDNA sequence represents a defective purified swine retrovirus
XX CC found in PK-15 cells containing the putative coding regions for viral
XX CC GAG, POL and ENV proteins. There are a few in frame stop codons and
XX CC apparent frame shifts in the given coding sequence which alter features
XX CC of the translation. This sequence and PCR fragments generated from the
XX CC sequence (see AAT74812-74882) could be used to screen organs for the
XX CC presence of porcine retroviruses prior to xenotransplantation.
XX CC Transplantation can increase the likelihood of retroviral activation if
XX CC intact and infectious proviruses are present. The porcine retroviral
XX CC sequence can be used to generate probes to determine the level (e.g.
XX CC copy number) of intact (i.e. potentially replicating) porcine provirus
XX CC sequences in a strain of xenograft transplantation donors. It can be
XX CC used to detect mutations, genetic lesions or viral recombinants and to
XX CC determine the histological localisation of activated retrovirus. Using
XX CC polymerase Chain Reaction DNA Quantitation (PQD) on blood mononuclear
XX CC cells, infectivity titration and susceptibility testing can be
XX CC performed. Ultimately animal donors without intact porcine retroviral
XX CC sequences or a lower copy number of viral elements could be selected.
XX
XX Sequence 7393 BP; 2004 A; 1801 C; 1896 G; 1692 T; 0 other;
SQ

```

Query Match 14.3%; Score 422; DB 18; Length 7393;
Best Local Similarity 99.8%; Pred. No. 5.3e-164;
Matches 472; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2082 gttagaagagcgctgaaggaagagagctgacccagggctgttgaagatgttcaa 2141
XX |||||||
DB 6470 gttagaagagcgctgaaggaagagagctgacccagggctgttgaagatgttcaa 6529
QY 2142 caggtctctcttgatgacccaccctgcttctgcttgcacggggccctagatcctgct 2201
XX |||||||
DB 6530 caggtctctcttgatgacccaccctgcttctgcttgcacggggccctagatcctgct 6589
QY 2202 cctgttactaacgttgggcttgccttaataatagttgttgccttctgttagaagc 2261
XX |||||||
DB 6590 cctgttactaacgttgggcttgccttaataatagttgttgccttctgttagaagc 6649
QY 2262 agtgaagcagctcagatcgttacttagcacaacgtacaaagggcttctgacaaag 2321
XX |||||||
DB 6650 agtgaagcagctcagatcgttacttagcacaacgtacaaagggcttctgacaaag 6709
QY 2322 agaaactgacctagccttccagttctaagattagaactataaacaagaagaagtg 2381
XX |||||||
DB 6710 agaaactgacctagccttccagttctaagattagaactataaacaagaagaagtg 6769
QY 2382 gggatgaagaagatgaataagcaacctaacctcccaagaagccaggaagtataaaga 2441
XX |||||||
DB 6770 gggatgaagaagatgaataagcaacctaacctcccaagaagccaggaagtataaaga 6829

Qy	208	gagacttgcaggttccacatcgccttccaaagtggagattcaattctatglttaagccacc	267
Db	2513	gagacttgcaggttccacatcgccttccaaagtggagattcaattctatglttaagccacc	2572

Db 8095 ggaataaaaatcctctgtgtgttgcatcaaaaaaa 8132

RESULT 10
AAC67007
ID AAC67007 standard; DNA; 633 BP.
XX
XX
AAC67007;
AC
XX
XX
D7 27-MAR-2001 (first entry)
XX
XX
DE PERV-MSL LTR SEQ ID NO: 7.
XX
XX
KW Xenotransplantation; infectious agent; vaccine; ds.
XX
XX
OS Porcine endogenous retrovirus.
XX
PN WO200071726-A1.
XX
PD 30-NOV-2000.
XX
PF 24-MAY-2000; 2000WO-US14296.
XX
PR 24-MAY-1999; 99US-0135631.
XX
PA (MAYO-) MAYO MEDICAL VENTURES.
XX
PI Federspiel MJ;
XX
DR WPI; 2001-032041/04.
XX
XX Inhibiting or preventing infectious agent transmission in mammalian
PT transplant recipients, by introducing recombinant DNA comprising DNA
PT encoding extracellular proteins of the agent into donor cells, such as
PT swine cells -
XX
XX Disclosure; Page 98; 144pp; English.
XX
XX The present invention provides a method to prevent the transmission of
CC infectious agents during xenotransplantation. This involves introducing
CC to donor swine cells a recombinant DNA encoding a peptide fragment from
CC the infectious agent, and then introducing these cells into the
CC transplant recipient.
XX
SQ Sequence 633 BP; 170 A; 125 C; 147 G; 191 T; 0 other;

Query Match 3.1%; Score 91; DB 22; Length 633;
Best Local Similarity 100.0%; Pred. No. 5.7e-28;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2846 cactcgaggcgagctctctaccctcgctgtgtacgactgtggcccgagcgctt 2905
|||||
DB 452 cactcgaggcgagctctctaccctcgctgtgtacgactgtggcccgagcgctt 511
|||||
QY 2906 ggaataaatacctctgtctgttgcacaa 2936
|||||
DB 512 ggaataaatacctctgtctgttgcacaa 542
|||||

RESULT 11
AAC67006
ID AAC67006 standard; DNA; 704 BP.
XX
XX
AAC67006;
AC
XX
XX
D7 27-MAR-2001 (first entry)
XX
XX
DE 1am1 LTR SEQ ID NO: 6.
XX
XX
KW Xenotransplantation; infectious agent; vaccine; ds.
XX
XX
OS Unidentified.
XX
PN WO200071726-A1.

PD 30-NOV-2000.
XX
XX
PF 24-MAY-2000; 2000WO-US14296.
XX
XX
PR 24-MAY-1999; 99US-0135631.
XX
XX
PA (MAYO-) MAYO MEDICAL VENTURES.
XX
XX
PI Federspiel MJ;
XX
DR WPI; 2001-032041/04.
XX
XX Inhibiting or preventing infectious agent transmission in mammalian
PT transplant recipients, by introducing recombinant DNA comprising DNA
PT encoding extracellular proteins of the agent into donor cells, such as
PT swine cells -
XX
XX Disclosure; Page 97-98; 144pp; English.
XX
XX The present invention provides a method to prevent the transmission of
CC infectious agents during xenotransplantation. This involves introducing
CC to donor swine cells a recombinant DNA encoding a peptide fragment from
CC the infectious agent, and then introducing these cells into the
CC transplant recipient.
XX
SQ Sequence 704 BP; 187 A; 162 C; 158 G; 197 T; 0 other;

Query Match 3.1%; Score 91; DB 22; Length 704;
Best Local Similarity 100.0%; Pred. No. 5.6e-28;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2846 cactcgaggcgagctctctaccctcgctgtgtacgactgtggcccgagcgctt 2905
|||||
DB 523 cactcgaggcgagctctctaccctcgctgtgtacgactgtggcccgagcgctt 582
|||||
QY 2906 ggaataaatacctctgtctgttgcacaa 2936
|||||
DB 583 ggaataaatacctctgtctgttgcacaa 613
|||||

RESULT 12
AAC67022
ID AAC67022 standard; DNA; 4918 BP.
XX
XX
AAC67022;
AC
XX
XX
D7 27-MAR-2001 (first entry)
XX
XX
DE PERV env protein coding sequence SEQ ID NO: 22.
XX
XX
KW Xenotransplantation; infectious agent; vaccine; ds.
XX
XX
OS Porcine endogenous retrovirus.
XX
PN WO200071726-A1.
XX
PD 30-NOV-2000.
XX
PF 24-MAY-2000; 2000WO-US14296.
XX
PR 24-MAY-1999; 99US-0135631.
XX
PA (MAYO-) MAYO MEDICAL VENTURES.
XX
XX
PI Federspiel MJ;
XX
DR WPI; 2001-032041/04.
XX
XX Inhibiting or preventing infectious agent transmission in mammalian
PT transplant recipients, by introducing recombinant DNA comprising DNA
PT encoding extracellular proteins of the agent into donor cells, such as
PT swine cells -

XX Claim 16; Page 109-111; 144pp; English.
PS
XX The present invention provides a method to prevent the transmission of
CC infectious agents during xenotransplantation. This involves introducing
CC to donor swine cells a recombinant DNA encoding a peptide fragment from
CC the infectious agent, and then introducing these cells into the
CC transplant recipient.
XX
SQ Sequence 4918 BP; 1382 A; 1173 C; 1183 G; 1180 T; 0 other;

Query Match 3.1%; Score 91; DB 22; Length 4918;
Best Local Similarity 100.0%; Pred. No. 4.3e-28;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2846 cactcggggccgagctcctctacccctcgtgtgtacagactgtggcccaagcgagctt 2905
|||||
DB 4736 cactcggggccgagctcctctacccctcgtgtgtacagactgtggcccaagcgagctt 4795
|||||
QY 2906 ggaataaaaatcctctgtctgttgcacaa 2936
|||||
DB 4796 ggaataaaaatcctctgtctgttgcacaa 4826
|||||

RESULT 13

AAC67021
ID AAC67021 standard; DNA: 6076 BP.

AC AAC67021;

DT 27-MAR-2001 (first entry)

DE PERV env protein coding sequence SEQ ID NO: 21.

KW Xenotransplantation; infectious agent; vaccine; ds.

OS Porcine endogenous retrovirus.

XX WO200071726-A1.

XX 30-NOV-2000.

XX 24-MAY-2000; 2000WO-US14296.

XX 24-MAY-1999; 99US-0135631.

XX (MAYO-) MAYO MEDICAL VENTURES.

XX Federspiel MJ;

XX WPI; 2001-032041/04.

XX Inhibiting or preventing infectious agent transmission in mammalian
PT transplant recipients, by introducing recombinant DNA comprising DNA
PT encoding extracellular proteins of the agent into donor cells, such as
PT swine cells -
XX
PS Claim 16; Page 107-109; 144pp; English.

CC The present invention provides a method to prevent the transmission of
CC infectious agents during xenotransplantation. This involves introducing
CC to donor swine cells a recombinant DNA encoding a peptide fragment from
CC the infectious agent, and then introducing these cells into the
CC transplant recipient.
XX

SQ Sequence 6076 BP; 1613 A; 1512 C; 1609 G; 1342 T; 0 other;

Query Match 3.1%; Score 91; DB 22; Length 6076;
Best Local Similarity 100.0%; Pred. No. 4.2e-28;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2846 cactcggggccgagctcctctacccctcgtgtgtacagactgtggcccaagcgagctt 2905
|||||
DB 572 cactcggggccgagctcctctacccctcgtgtgtacagactgtggcccaagcgagctt 631
|||||
QY 2906 ggaataaaaatcctctgtctgttgcacaa 2936
|||||
DB 632 ggaataaaaatcctctgtctgttgcacaa 662
|||||

RESULT 14

AAC67032
ID AAC67032 standard; DNA: 6076 BP.

AC AAC67032;

DT 27-MAR-2001 (first entry)

DE PERV env protein coding sequence SEQ ID NO: 32.

KW Xenotransplantation; infectious agent; vaccine; ds.

OS Porcine endogenous retrovirus.

XX WO200071726-A1.

XX 30-NOV-2000.

XX 24-MAY-2000; 2000WO-US14296.

XX 24-MAY-1999; 99US-0135631.

XX (MAYO-) MAYO MEDICAL VENTURES.

XX Federspiel MJ;

XX WPI; 2001-032041/04.

XX Inhibiting or preventing infectious agent transmission in mammalian
PT transplant recipients, by introducing recombinant DNA comprising DNA
PT encoding extracellular proteins of the agent into donor cells, such as
PT swine cells -
XX
PS Claim 16; Page 117-119; 144pp; English.

CC The present invention provides a method to prevent the transmission of
CC infectious agents during xenotransplantation. This involves introducing
CC to donor swine cells a recombinant DNA encoding a peptide fragment from
CC the infectious agent, and then introducing these cells into the
CC transplant recipient.
XX

SQ Sequence 6076 BP; 1619 A; 1505 C; 1598 G; 1354 T; 0 other;

Query Match 3.1%; Score 91; DB 22; Length 6076;
Best Local Similarity 100.0%; Pred. No. 4.2e-28;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2846 cactcggggccgagctcctctacccctcgtgtgtacagactgtggcccaagcgagctt 2905
|||||
DB 572 cactcggggccgagctcctctacccctcgtgtgtacagactgtggcccaagcgagctt 631
|||||
QY 2906 ggaataaaaatcctctgtctgttgcacaa 2936
|||||
DB 632 ggaataaaaatcctctgtctgttgcacaa 662
|||||

RESULT 15

AAC67019
ID AAC67019 standard; DNA: 7362 BP.

AC AAC67019;

DT 27-MAR-2001 (first entry)

```

XX PERV env protein coding sequence SEQ ID NO: 19.
DE
XX Xenotransplantation; infectious agent; vaccine; ds.
KW
XX Porcine endogenous retrovirus.
OS
XX WO200071726-A1.
PN
XX 30-NOV-2000.
PD
XX 24-MAY-2000; 2000WO-US14296.
PF
XX 24-MAY-1999; 99US-0135631.
PR
XX (MAYO-) MAYO MEDICAL VENTURES.
PA
XX Federspiel MJ;
PI
XX WPI; 2001-032041/04.
DR
XX
XX Inhibiting or preventing infectious agent transmission in mammalian
PT transplant recipients, by introducing recombinant DNA comprising DNA
PT encoding extracellular proteins of the agent into donor cells, such as
PT swine cells -
XX
XX Claim 16; Page 101-104; 14pp; English.
PS
XX
XX The present invention provides a method to prevent the transmission of
CC infectious agents during xenotransplantation. This involves introducing
CC to donor swine cells a recombinant DNA encoding a peptide fragment from
CC the infectious agent, and then introducing these cells into the
CC transplant recipient.
XX
XX Sequence 7362 BP; 1997 A; 1821 C; 1881 G; 1663 T; 0 other;
SQ

```

Query Match 3.1%; Score 91; DB 22; Length 7362;
 Best Local Similarity 100.0%; Pred No. 4.1e-28;
 Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 2846 cactcgggggccgagctctaccctcgtggtgtaagactgtggcccccagcgctt 2905
   |||||||
DB 637 cactcgggggccgagctctaccctcgtggtgtaagactgtggcccccagcgctt 696
QY 2906 ggaataaaatcctctgtgtgttgcataa 2936
   |||||||
DB 697 ggaataaaatcctctgtgttgcataa 727

```

Search completed: February 24, 2002, 07:43:57
 Job time: 16942 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 23, 2002, 21:40:34 ; Search time 624.99 Seconds

(without alignments)
4554.185 Million cell updates/sec

Title: US-09-171-553B-1

Perfect score: 3320

Sequence: 1 gaattcgagcgccgctgcac.....aaaaaaaaaaaaaaaaa 3320

Scoring table: IDENTITY NUC

Searched: 930621 seqs, 428662619 residues 1861242

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Database : N_Geneseq_1101.*

1: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3320	100.0	3320	18	AAV09698 Porcine retrovirus
2	3179	95.8	4402	22	AAV09698 PERV env protein c
3	3175.8	95.7	7362	22	AAV09698 PERV env protein c
4	2976.4	89.7	7873	22	AAV09698 PERV env protein c
5	2916.8	87.9	8196	18	AAV09698 Porcine retrovirus
6	2916	87.8	8209	18	AAV09698 Porcine retrovirus
7	2915.2	87.8	7892	18	AAV09698 Porcine retrovirus
8	2915.2	87.8	8132	18	AAV09698 Porcine retrovirus
9	2721.4	82.0	4918	22	AAV09698 PERV env protein c
10	2420.6	72.9	8060	18	AAV09698 Porcine retrovirus
11	2420.6	72.9	8060	22	AAV09698 Porcine retrovirus

12	2415.6	72.8	6076	22	AAV09698
13	2414	72.7	6076	22	AAV09698
14	1771.8	53.4	7333	22	AAV09698
15	1771.8	53.4	7333	22	AAV09698
16	1175.8	35.4	8655	20	AAV09698
17	1131	34.1	3482	20	AAV09698
18	1088	32.8	8088	21	AAV09698
19	1088	32.8	8088	21	AAV09698
20	941.4	28.4	8278	22	AAV09698
21	940	28.3	8892	16	AAV09698
22	933.6	28.1	8323	16	AAV09698
23	933.6	28.1	8323	16	AAV09698
24	933.6	28.1	8323	16	AAV09698
25	932	28.1	8323	15	AAV09698
26	932	28.1	10367	15	AAV09698
27	932	28.1	10367	15	AAV09698
28	932	28.1	10367	15	AAV09698
29	916.2	27.6	3643	21	AAV09698
30	916.2	27.6	5178	21	AAV09698
31	916.2	27.6	7308	18	AAV09698
32	916.2	27.6	7308	18	AAV09698
33	916.2	27.6	7616	18	AAV09698
34	916.2	27.6	8330	20	AAV09698
35	916.2	27.6	8332	17	AAV09698
36	916.2	27.6	8332	18	AAV09698
37	916.2	27.6	8332	19	AAV09698
38	916.2	27.6	8332	19	AAV09698
39	916.2	27.6	8332	19	AAV09698
40	916.2	27.6	8332	20	AAV09698
41	916.2	27.6	8332	21	AAV09698
42	916.2	27.6	8332	21	AAV09698
43	916.2	27.6	8332	21	AAV09698
44	916.2	27.6	12335	18	AAV09698
45	916.2	27.6	36538	18	AAV09698

ALIGNMENTS

AAV09698	1	AAV09698 standard; DNA: 3320 BP.
AC	AAV09698;	
XX	19-MAY-1998 (first entry)	
DE	Porcine retrovirus pol and env DNA.	
XX	Porcine retrovirus; POEV; POL protein; ENV protein; vaccine;	
KW	diagnosis; xenotransplantation; prophylactic; therapeutic; ds.	
XX	Porcine retrovirus.	
OS	Porcine retrovirus.	
XX		
FT	Key	Location/Qualifiers
FT	CDS	23..2793
FT		/*tag= a
FT		/product= POL protein
FT		/note= "polymerase protein"
FT	CDS	2642..3297
FT		/*tag= b
FT		/product= ENV protein
FT		/note= "envelope protein"
XX	W09740167-A1.	
PN	30-OCT-1997.	
XX		
PD	18-APR-1997;	97W0-GB01087.
XX		
PF	10-FEB-1997;	97GB-0002668.
XX	19-APR-1996;	96GB-0008164.
PR		
XX		

PA (IMUT-) IMUTRAN LTD.
PA (QONE-) Q-ONE BIOTECH LTD.
XX Galbraith DN, Haworth C, Lees GM, Smith KT.
PI WPI, 1997-535851/49.
XX
DR
XX
PT Polynucleotide encoding porcine retrovirus expression product -
PT useful to develop products for use in vaccines, diagnosis and
PT xeno-transplantation
XX
PS Claim 4; Fig 1; 69pp; English.
CC This DNA sequence encodes the porcine retrovirus (PoEV) polymerase (POL)
CC and envelope (ENV) proteins. These proteins can be used to develop viral
CC vaccines, antisense nucleic acids, ribozymes and other antiviral agents.
CC They can also be used in xeno-transplantation technology and as
CC diagnostic tools.
XX
SQ Sequence 3320 BP; 922 A; 844 C; 832 G; 722 T; 0 other;

Query Match 100.0%; Score 3320; DB 18; Length 3320;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gaattcgccgagcgttcagatgctcttctgctgagatatacccccactagccaa 60
Db 1 gaattcgccgagcgttcagatgctcttctgctgagatatacccccactagccaa 60
QY 61 ccaatttttgcttcgaatggagatccaggtacgggaagaaacggcgacactctg 120
Db 61 ccaatttttgcttcgaatggagatccaggtacgggaagaaacggcgacactctg 120
QY 121 accgactgccccaaaggttcaagaactcccccagcaactcttgaagaaacccatacaga 180
Db 121 accgactgccccaaaggttcaagaactcccccagcaactcttgaagaaacccatacaga 180
QY 181 gaactgccaactcagatccaaacacccctcaagtgacccctctcagtaactgtagtac 240
Db 181 gaactgccaactcagatccaaacacccctcaagtgacccctctcagtaactgtagtac 240
QY 241 ctgcttctgcccgaagcccaaacagagactgttgaagatgacgaagacactacgtc 300
Db 241 ctgcttctgcccgaagcccaaacagagactgttgaagatgacgaagacactacgtc 300
QY 301 gaattgctgaactcagatccaaacacccctcaagtgacccctctcagtaactgtagtac 360
Db 301 gaattgctgaactcagatccaaacacccctcaagtgacccctctcagtaactgtagtac 360
QY 361 gtaacatactggggtacagtttgcgggcccgaagcgtatgctgacgagacgagaa 420
Db 361 gtaacatactggggtacagtttgcgggcccgaagcgtatgctgacgagacgagaa 420
QY 421 aaactgtagtccagatccaaacacccctcaagtgacccctctcagtaactgtagtac 480
Db 421 aaactgtagtccagatccaaacacccctcaagtgacccctctcagtaactgtagtac 480
QY 481 acagctgtagtccagatccaaacacccctcaagtgacccctctcagtaactgtagtac 540
Db 481 acagctgtagtccagatccaaacacccctcaagtgacccctctcagtaactgtagtac 540
QY 541 ccgctaaacaaagaaagggggtatctcctggcctcctgagcaccagaagcattgatt 600
Db 541 ccgctaaacaaagaaagggggtatctcctggcctcctgagcaccagaagcattgatt 600
QY 601 gctatacaaaagcccgctgagacgacacgctcctgagcaccagaagcattgatt 660
Db 601 gctatacaaaagcccgctgagacgacacgctcctgagcaccagaagcattgatt 660
QY 661 tttaacctatgtgagtagcgttaagggatgagcccgagagatttaaccacaaacctta 720
Db 661 tttaacctatgtgagtagcgttaagggatgagcccgagagatttaaccacaaacctta 720

QY 721 ggaacataagagagacgttgcctacactgtcaaaagcgttgcctgtagccagtgct 780
Db 721 ggaacataagagagacgttgcctacactgtcaaaagcgttgcctgtagccagtgct 780
QY 781 tggcccgatgtctgaagctatcgacgttgcgacatactgtgttaagagcgttaca 840
Db 781 tggcccgatgtctgaagctatcgacgttgcgacatactgtgttaagagcgttaca 840
QY 841 ttgacttggagaagataaactgttaagcccccatgcatgtgaataactgttcg 900
Db 841 ttgacttggagaagataaactgttaagcccccatgcatgtgaataactgttcg 900
QY 901 cagccccagacagatgtagtgaacaaacgcccagatgacccactatcaaacgtctctc 960
Db 901 cagccccagacagatgtagtgaacaaacgcccagatgacccactatcaaacgtctctc 960
QY 961 acagagagagtgacttgcctcaccagcgtctcaacccctgcacacttctgctgaa 1020
Db 961 acagagagagtgacttgcctcaccagcgtctcaacccctgcacacttctgctgaa 1020
QY 1021 ggaactgtagaacacagtgatcatgtatgccaactatgtatgagagagctggc 1080
Db 1021 ggaactgtagaacacagtgatcatgtatgccaactatgtatgagagagctggc 1080
QY 1081 cgaagagaccttacaagatacagcgttgcgagagagtgtaacctgttcaactgag 1140
Db 1081 cgaagagaccttacaagatacagcgttgcgagagagtgtaacctgttcaactgag 1140
QY 1141 agcagctatgtgtggaaggtgaagagtgctgtggggcgagtggtgagggagccgc 1200
Db 1141 agcagctatgtgtggaaggtgaagagtgctgtggggcgagtggtgagggagccgc 1200
QY 1201 acgactgtagcagacagcctgcggaaggaacttaagcgcaaaagcgtgactatgccc 1260
Db 1201 acgactgtagcagacagcctgcggaaggaacttaagcgcaaaagcgtgactatgccc 1260
QY 1261 ctcaacgaagcttgcggtgcggaaggaatccataaactatatacgaagacagag 1320
Db 1261 ctcaacgaagcttgcggtgcggaaggaatccataaactatatacgaagacagag 1320
QY 1321 tatgcttgcgacgtgcagacgttgcggaaggaatccataaactatatacgaagacagag 1380
Db 1321 tatgcttgcgacgtgcagacgttgcggaaggaatccataaactatatacgaagacagag 1380
QY 1381 tcaagcagagaggaataaagacaaagaggaattcctaagcctatagaagccttaccat 1440
Db 1381 tcaagcagagagaggaataaagacaaagaggaattcctaagcctatagaagccttaccat 1440
QY 1441 ttgccaanaaggtgactatatacactgtccctgagacatcagaagccaaagatccata 1500
Db 1441 ttgccaanaaggtgactatatacactgtccctgagacatcagaagccaaagatccata 1500
QY 1501 tctaaaggaacagatggtgcgacgggttgcgaagcagagccagggcgtttaacct 1560
Db 1501 tctaaaggaacagatggtgcgacgggttgcgaagcagagccagggcgtttaacct 1560
QY 1561 ctgctataatagaacagcccaagcccaagcccaagcccaagcccaagcccaagcccaag 1620
Db 1561 ctgctataatagaacagcccaagcccaagcccaagcccaagcccaagcccaagcccaag 1620
QY 1621 tggcagaagataaataagacacagcttctcgtgagaccccgagggagacctgtatacc 1680
Db 1621 tggcagaagataaataagacacagcttctcgtgagaccccgagggagacctgtatacc 1680
QY 1681 tcatatggagaaagtctcctcccaaaagaggttagaatgttccaaacatacat 1740
Db 1681 tcatatggagaaagtctcctcccaaaagaggttagaatgttccaaacatacat 1740
QY 1741 cgtctaacccactagagactaaacacccctgagcaggttgcagaaacatccctcatcat 1800
Db 1741 cgtctaacccactagagactaaacacccctgagcaggttgcagaaacatccctcatcat 1800

QY 2122 ccagcttcgttgcgccagtaagtcaggaactggccaaagatatgtggatcttgtaaa 2181
 Db 3134 ccagcttcgttgcgccagtaagtcaggaactggccaaagatatgtggatcttgtaaa 3193
 QY 2182 ctgcatgtgtacacagaccacaaagtcagagcaggtagagagatgaatagaacatt 2241
 Db 3194 ctgcatgtgtacacagaccacaaagtcagagcaggtagagagatgaatagaacatt 3253
 QY 2242 aaagaagcccttaccacaaatggacacagagactggactatgatgttgatgctctctg 2301
 Db 3254 aaagaagcccttaccacaaatggacacagagactggactatgatgttgatgctctctg 3313
 QY 2302 ccccttgccttctttagagtgaggaacacccctggacagcttggcctgacccctatgaa 2361
 Db 3314 ccccttgccttctttagagtgaggaacacccctggacagcttggcctgacccctatgaa 3373
 QY 2362 ttgctctacggggagaccccccgttggcagaaattgctcttcacatagtgctgattg 2421
 Db 3374 ttactctacggggagaccccccgttggcagaaattgctcttcacatagtgctgattg 3433
 QY 2422 ctgcttccagccttggcttctctaggtctaaagcgctgagtggtggagcagcgagcg 2481
 Db 3434 ctgcttccagccttggcttctctaggtctaaagcgctgagtggtggagcagcgagcg 3493
 QY 2482 tggaaagagctccggagagcctactc--aggaagagacttgcaagttccacacatgcctc 2538
 Db 3494 tggagagagctccggagagcctactcagggagggagagacttgcaagttccacacatgcctc 3553
 QY 2539 caagttgagagattcagctctctgttagagcgccacgctggcagaaacctcgagactcggtg 2598
 Db 3554 caagttgagagattcagctctctgttagagcgccacgctggcagaaacctcgagactcggtg 3613
 QY 2599 aaggagacattatctgcgacttttgacacacacaaagcgctgtgaagtctgaagaaatcccc 2658
 Db 3614 aaggagacattatctgcgacttttgacacacacaaagcgctgtgaagtctgaagaaatcccc 3673
 QY 2659 acctggtatcatgcatccacgcttaagcgcgacacactccgattcgggggtgaaagcc 2718
 Db 3674 acctggtatcatgcatccacgcttaagcgcgacacactccgattcgggggtgaaagcc 3733
 QY 2719 gaaaagctgaaaatcccccttaagcttcgctccatccatcggtgttcccttaactctgcaat 2778
 Db 3734 gaaaagctgaaaatcccccttaagcttcgctccatccatcggtgttcccttaactctgcaat 3793
 QY 2779 aactcccaagtttaatgtaaacgcttggcagcgaccccaactccatcaaacctctatc 2838
 Db 3794 aactcccaagtttaatgtaaacgcttggcagcgaccccaactccatcaaacctctatc 3853
 QY 2839 tctcaactggttactactgactcgggtacaggtatataatataaacagcactcaagggga 2898
 Db 3854 tctcaactggttactactgactcgggtacaggtatataatataaacagcactcaagggga 3913
 QY 2899 ggtctcccttgggagactggtggcgtgaatatatgtctccttcgtacagtaactccctg 2958
 Db 3914 ggtctcccttgggagactggtggcgtgaatatatgtctccttcgtacagtaactccctg 3973
 QY 2959 tcccaatgaccagggacacaccccgatctactcgttcttcagggttttaagtttgccc 3018
 Db 3974 tcccaatgaccagggacacaccccgatctactcgttcttcagggttttaagtttgccc 4033
 QY 3019 aggaaccccccaataatgaagaatatgtgtgaaatctctcaggaattcctcttgcaagaaag 3078
 Db 4034 aggaaccccccaataatgaagaatatgtgtgaaatctctcaggaattcctcttgcaagaaag 4093
 QY 3079 gaagctgataacttctaaatgagtggaattggaatggcagctctcagcaagaagagct 3138
 Db 4094 gaagctgataacttctaaatgagtggaattggaatggcagctctcagcaagaagagct 4153
 QY 3139 aagttactcttctgttaacaatcctacaggttaataatcaatgaatgaagcagggg 3198
 Db 4154 aagttactcttctgttaacaatcctacaggttaataatgaatgaagcagggg 4213

QY 3199 atggaaagattggcaacacggttacaacaaagatgttacaagaataagcaataagcttcca 3258
 Db 4214 atggaaagattggcaacacggttacaacaaagatgttacaagaataagcaataagcttcca 4273
 QY 3259 ttgcgttagacctagattacttaaaaaaagtttccactaaataaaaaaagaaataaa 3317
 Db 4274 ttgcgttagacctagattacttaaaaaaagtttccactgaaataaagaaacaaagaaata 4332

RESULT 3

AAC67019
 ID AAC67019 standard; DNA; 7362 BP.

AC AAC67019;

DT 27-MAR-2001 (first entry)

DE PERV env protein coding sequence SEQ ID NO: 19.

KW Xenotransplantation; infectious agent; vaccine; ds.

OS Porcine endogenous retrovirus.

PN WO200071726-A1.

PD 30-NOV-2000.

PF 24-MAY-2000; 2000WO-US14296.

PR 24-MAY-1999; 99US-013631.

PA (MAYO-) MAYO MEDICAL VENTURES.

PI Federspiel MJ.

DR WPI: 2001-032041/04.

XX

PT Inhibiting or preventing infectious agent transmission in mammalian

PT transplant recipients, by introducing recombinant DNA comprising DNA

PT encoding extracellular proteins of the agent into donor cells, such as

PT swine cells -

PS Claim 16; Page 101-104; 144pp; English.

CC The present invention provides a method to prevent the transmission of

CC infectious agents during xenotransplantation. This involves introducing

CC to donor swine cells a recombinant DNA encoding a peptide fragment from

CC the infectious agent, and then introducing these cells into the

CC transplant recipient.

CC

XX

SO

Sequence 7362 BP; 1997 A; 1821 C; 1881 G; 1663 T; 0 other;

Query Match 95.7%; Score 3175.8; DB 22; Length 7362;

Best Local Similarity 97.9%; Pred. No. 0;

Matches 3229; Conservative 0; Mismatches 67; Indels 3; Gaps 1;

QY 22 gatgctctctctgctgctgagattaaacccacatagcaaacacttttggcttgaatgg 81
 Db 3633 gatgctctctctgctgctgagattaaacccacatagcaaacacttttggcttgaatgg 3692
 QY 82 aagatccaggtacgaggaagaacggcgagctcacttgagccgactggcccaagggttc 141
 Db 3693 aagatccaggtacgaggaagaacggcgagctcacttgagccgactggcccaagggttc 3752
 QY 142 aagaactcccgacacatcttgcagaaagccctacacagagaaactggccaacttcagatc 201
 Db 3753 aagaactcccgacacatcttgcagaaagccctacacagagaaactggccaacttcagatc 3812
 QY 202 caaacccctcagggagacccctctccagtaagtgatgatacctgctctggcgggagcacc 261
 Db 3813 caaacccctcagggagacccctctccagtaagtgatgatacctgctctggcgggagcacc 3872

QY 262 aaacagactgtctagaaggtacgaagcactctgtgaaattgtctgactaggctac 321
|||||
Db 3873 aaacagactgtctagaaggtacgaagcactctgtgaaattgtctgactaggctac 3932
QY 322 agagcctctgtagaagggccagatttctgagagagaggttaactacttgggtacgt 381
|||||
Db 3933 agagcctctgtagaagggccagatttctgagagagaggttaactacttgggtacgt 3992
QY 382 ttggtgggcgggcagcagatgtgacgagcagcagaaactgagtcagataccg 441
|||||
Db 3993 ttggtgggcgggcagcagatgtgacgagcagcagaaactgagtcagataccg 4052
QY 442 gccccaacacgaacgaacaaagtgtgagagtttttgggagagctgatttgcagactg 501
|||||
Db 4053 gccccaacacgaacgaacaaagtgtgagagtttttgggagagctgatttgcagactg 4112
QY 502 tggatcccggggttttgagaccttagcagcccaactacccgcttaacaaagaaagg 561
|||||
Db 4113 tggatcccggggttttgagaccttagcagcccaactacccgcttaacaaagaaagg 4172
QY 562 ggaattccctgggctccctgagcacaagaagcatttgatgtatcaaaaaggccctgtg 621
|||||
Db 4173 ggaattccctgggctccctgagcacaagaagcatttgatgtatcaaaaaggccctgtg 4232
QY 622 agcgcaacctgtctgtccctccctgtgacgttaactaaaccttacccttattgtgagtg 681
|||||
Db 4233 agcgcaacctgtctgtccctccctgtgacgttaactaaaccttacccttattgtgagtg 4292
QY 682 cgttaagagtagcccgagaggttttaaccacacacttagacacagagagagactgtt 741
|||||
Db 4293 cgttaagagtagcccgagaggttttaaccacacacttagacacagagagagactgtt 4352
QY 742 gctcaactgtcaaaagagctgtactctgttagcagatgtgtgcccgtatgtctgaagct 801
|||||
Db 4353 gctcaactgtcaaaagagctgtactctgttagcagatgtgtgcccgtatgtctgaagct 4412
QY 802 atcgcaactgtgcccactactggtcaagaagcagcagcaaatgtgactttggagacagata 861
|||||
Db 4413 atcgcaactgtgcccactactggtcaagaagcagcagcaaatgtgactttggagacagata 4472
QY 862 actgttaataagccccactgtcatgtgagagacatcgttcggagagccccagacagatgagt 921
|||||
Db 4473 actgttaataagccccactgtcatgtgagagacatcgttcggagagccccagacagatgagt 4532
QY 922 accaagcccgcatgacccactatcaaaagcctgtctctcaagaagaggtcacttcgct 981
|||||
Db 4533 accaagcccgcatgacccactatcaaaagcctgtctctcaagaagaggtcacttcgct 4592
QY 982 ccacacgcgctctcaacactgcaactctctgctgaagaagacatgataccagtgact 1041
|||||
Db 4593 ccacacgcgctctcaacactgcaactctctgctgaagaagacatgataccagtgact 4652
QY 1042 catgattgcaacaactatgtattgagagagctgggtcccgcaagagaccttacagacata 1101
|||||
Db 4653 catgattgcaacaactatgtattgagagagctgggtcccgcaagagaccttacagacata 4712
QY 1102 ccgctgactgtgagaagtgctaaactgttactacgacggaagcagctatgtgtggaagt 1161
|||||
Db 4713 ccgctgactgtgagaagtgctaaactgttactacgacggaagcagctatgtgtggaagt 4772
QY 1162 aagaagatgtgctgggagcagtggtgagcagagcccgacgatactcgggccaacaacctg 1221
|||||
Db 4773 aagaagatgtgctgggagcagtggtgagcagagcccgacgatactcgggccaacaacctg 4832
QY 1222 ccggaaggaactcagcagcaaaagctgagctcaatgcccctcaacgaacttgcgctg 1281
|||||
Db 4833 ccggaaggaactcagcagcaaaagctgagctcaatgcccctcaacgaacttgcgctg 4892
QY 1282 gccgaaggaatccataaactttatcagacagcagatgataccttctgactgtgacac 1341
|||||
Db 4893 gccgaaggaatccataaactttatcagacagcagatgataccttctgactgtgacac 4952
QY 1342 gtacacggygcattctataaacaaggggttcttaacctcagcagggaggaataaag 1401

Db 4953 gtacacggygcattctataaacaaggggttcttaacctcagcagggaggaataaag 5012
QY 1402 aacaagaagaatctctaagcctatagaagccttaacttggccaanaagctgactat 1461
|||||
Db 5013 aacaagaagaatctctaagcctatagaagccttaacttggccaanaagctgactat 5072
QY 1462 atacactgtctctgacatcagaagaacgaagatccatcatatcagaagggagacagtgct 1521
|||||
Db 5073 atacactgtctctgacatcagaagaacgaagatccatcatatcagaagggagacagtgct 5132
QY 1522 gacccggttgcacagcagcagccagcctgtttaaacttctgactataatagaacgcc 1581
|||||
Db 5133 gacccggttgcacagcagcagccagcctgtttaaacttctgactataatagaacgcc 5192
QY 1582 aagccccaggaacccagacagcagcagcagcagcagcagcagcagcagcagcagc 1641
|||||
Db 5193 aagccccaggaacccgagacagcagcagcagcagcagcagcagcagcagcagcagc 5252
QY 1642 gacccagttctgagactccggaagggagcctgtctatccatatagtggaagaaactcctg 1701
|||||
Db 5253 gacccagttctgagactccggaagggagcctgtctatccatatagtggaagaaactcctg 5312
QY 1702 ccccaaaaagaggggttgaataatgtccaagaagataactgtctaaacccactggaact 1761
|||||
Db 5313 ccccaaaaagaggggttgaataatgtccaagaagataactgtctaaacccactggaact 5372
QY 1762 aacacccctcaagattgtgcaagaacatcccttaactatgtctgaagctcaggaagt 1821
|||||
Db 5373 aacacccctcaagattgtgcaagaacatcccttaactatgtctgaagctcaggaagt 5432
QY 1822 gctgaactcgggtgtcaaacatgtgtgctccctgcaagctgtgttaatgtcaactccaga 1881
|||||
Db 5433 gctgaactcgggtgtcaaacatgtgtgctccctgcaagctgtgttaatgtcaactccaga 5492
QY 1882 atacctccaggaagaagactaaggggaagccacacagcgcctcacttgggaagtgaactc 1941
|||||
Db 5493 atacctccaggaagaagactaaggggaagccacacagcgcctcacttgggaagtgaactc 5552
QY 1942 actgaggtaaagccggtcaaatacggaacaaatactatgtgttttgaagacactt 2001
|||||
Db 5553 actgaggtaaagccggtcaaatacggaacaaatactatgtgttttgaagacactt 5612
QY 2002 taagatgtgtgaagcttatctactaagaagaagagacttaaacogtgggtgagaa 2061
|||||
Db 5613 taagatgtgtgaagcttatctactaagaagaagagacttaaacogtgggtgagaa 5672
QY 2062 atactggaagaaatcttccaagatttgaataactaaggttaatgagtcagacatgct 2121
|||||
Db 5673 atactggaagaaatcttccaagatttgaataactaaggttaatgagtcagacatgct 5732
QY 2122 ccagcttctgttgcacagtagtaagcagcagcagcagcagcagcagcagcagcagc 2181
|||||
Db 5733 ccagcttctgttgcacagtagtaagcagcagcagcagcagcagcagcagcagcagc 5792
QY 2182 cftgacttgcataagaaccccaagcctcagaagcagcagcagcagcagcagcagcagc 2241
|||||
Db 5793 cftgacttgcataagaaccccaagcctcagaagcagcagcagcagcagcagcagcagc 5852
QY 2242 aagaagacccctatacaaatgtgacacagagcctggaacttaatgtgattgactcctg 2301
|||||
Db 5853 aagaagacccctatacaaatgtgacacagagcctggaacttaatgtgattgactcctg 5912
QY 2302 ccccttgtgcttttaaggtgaggaacacccctggaacagtttgggtcgaacccctatgaa 2361
|||||
Db 5913 ccccttgtgacttttaaggtttaggaacacccctggaacagtttgggtcgaacccctatgaa 5972
QY 2362 ttgcttaagggggagccccccttggcgaagaatttgccttgcacagatgagcagtg 2421
|||||
Db 5973 ttaccttaagggggagccccccttggcgaagaatttgccttgcacagatgagcagtg 6032
QY 2422 ctgacttccagccttctctctagcacaagcgcctcaggtgggtgagcagcagcag 2481
|||||

Db	6033	ctgcttcccgagccttctgttctctcctaggtcctaaagcacttgagtggttgagaaacagagcg	6092
QY	2482	tggaaagcagctcccgagagcgactactc---agagagagacttgcgaagttccacatcgcttc	2538
Db	6093	tggagagcaacatcccgagagagcctactcagagagagagagacttgcagatcccaatcgtttc	6152
QY	2539	caagtttgaagatltcagctcctatgttttaagacgcacacgcgagagagaaacctcgagatctcggtg	2598
Db	6153	caagtgaggagatctcagctctcagcttctagacgcgcacgcgtgcagagaaacctcgagatctcggtg	6212
QY	2599	aagagaccttactctcgtacttttgcacacaccaaagcgctgtgaaagtctgaagaaatccccc	2658
Db	6213	aagggcccttctccagctacttttgcacacaccaaagcgctgtgaaagtctgaagaaatccccc	6272
QY	2659	aacctggtatccatgtatccccaactgaagccggcgccacacctcccgatctcggtgtgaaagcc	2718
Db	6273	aacctggtatcccatgtatccccaactgaagccggcgccacacctcccgatctcggtgtgaaagcc	6322
QY	2719	gaaagaagacttgaaaatcccccctaagacttgcgctccatcagctgtgttctccttaactcgtcat	2787
Db	6333	gaaagaagacttgaaaatcccccctaagacttgcgctccatcagctgtgttctccttaactcgtcat	6392
QY	2779	aacctctcaagtttaatggttaaacgcgcttgttgcagacgcccgaacctcccaataaccttacc	2838
Db	6393	aacctctcaagtttaatggttaaacgcgcttgttgcagacgcccgaacctcccaataaccttacc	6452
QY	2839	tctcaactcgttlaacttactctaacctccggtacagatattatataacagacatccaaagggga	2898
Db	6453	tctcaactcgttlaacttactctaacctccggtacagatattatataacagacatccaaagggga	6512
QY	2899	ggctccctctggggagacctgtgtgcgctcgaattatatagtctgccttcgatatcagtaatccctgg	2958
Db	6513	ggctccctctggggagacctgtgtgcgctcgaattatatagtctgccttcgatatcagtaatccctgg	6572
QY	2959	tctcaatgagcaagcgacacaccccccgatattactccgttcgttaacgggtttacgtttgcgcc	3018
Db	6573	tctcaatgagcaagcgacacaccccccgatattactccgttcgttaacgggtttacgtttgcgcc	6632
QY	3019	agagaccccccaaatgaagaagaatattgttgcgaatccacagatattcctcttgcagaacatg	3078
Db	6633	agagaccccccaaatgaagaagaatattgttgcgaatccacagatattcctcttgcagaacatg	6692
QY	3079	gagctgtgataactcttcaatggaatgtggaattgggaattggccagctctctcgaagacaagat	3138
Db	6693	gagctgtgataactcttcaatggaatgtggaattgggaattggccagctctctcgaagacaagat	6752
QY	3139	aagttactactcttgtttaaacaatcccaacggtataatcaatttaattatrttgcctatggag	3198
Db	6753	aagttactactcttgtttaaacaatcccaacggtataatcaatttaattatrttgcctatggag	6812
QY	3199	atgggaagaattggaacagcggtacaaaaagaatgtaacgaaataagcaaaataagctgtca	3258
Db	6813	atgggaagaattggaacagcggtacaaaaagaatgtaacgaaataagcaaaataagctgtca	6872
QY	3259	tctcgttagactcagatattcttaaaaaatagtttcctccttaaaaaaataaaaaaataaaaaa	3317
Db	6873	tctcgttagactcagatattcttaaaaaatagtttcctcctccttaaaaaaataaaaaaataaaaaa	6931

RESULT	4
AAC67023	
ID	AAC67023 standard; DNA; 7873 BP
XX	

DT 27-MAR-2001 (first entry)
XX

DE PERV env protein coding sequence SEQ ID NO: 23.

Xenotransplantation; infectious agent; vaccine; ds

OS Porcine endogenous retrovirus.

XX

XX	WO200071726-A1.
PN	
PD	30-NOV-2000.
XX	
PF	24-MAY-2000; 2000WO-US14296.
XX	
PR	24-MAY-1999; 99US-0135631.
XX	
PA	(MAYO-) MAYO MEDICAL VENTURES.
XX	
PI	Federspiel MJ;
XX	
DR	WPI: 2001-032041/04.
XX	
PT	Inhibiting or preventing infectious agent transmission in mammalian
PT	transplant recipients, by introducing recombinant DNA comprising DNA
PT	encoding extracellular proteins of the agent into donor cells, such as
PT	swine cells -
XX	
PS	Claim 16; Page 112-115; 14pp; English.
XX	
CC	The present invention provides a method to prevent the transmission of
CC	infectious agents during xenotransplantation. This involves introduc
CC	to donor swine cells a recombinant DNA encoding a peptide fragment fro
CC	the infectious agent, and then introducing these cells into the
CC	transplant recipient.
XX	
SO	Sequence 7873 BP; 2200 A; 1913 C; 1962 G; 1798 T; 0 other;

Query Match	89.78;	Score 2976.4;	DB 22;	Length 7873;
Best Local Similarity	95.1%;	Pred. No. 0;		
Matches 3139;	Conservative	0;	Mismatches 71;	Indels 90;
				Gaps 3

QY	22	gaagcctctcttcgctgcgagatlaaacacccactagccaaccacttlttgcttcgaaatg	81
Db	2447	gatgcctcttcttcgcgcgagatlaaacacccactagccaaccacttlttgcttcgaaatg	250
QY	82	agagatccaagttaagggaaagaacccgggcaagctcaacccctggagcccgactggcccaaggttc	141
Db	2507	agagatccaagttaaagggaagaacccgggcaagctcaacccctggagcccgactggcccaaggttc	256
QY	142	aagaactcccccgaaccatctttgacaaagccctacaacaaggagaccttggccaactcaagatc	201
Db	2567	aagaactcccccgaaccatctttgacaaagccctacaacaaggagaccttggccaactcaagatc	262
QY	202	caacaacctcaagtgatgacccctcttcagtaatgataagactgctcttgccggagccacc	261
Db	2627	caacaacctcaagtgatgacccctcttcagtaatgataagactgctcttgccggagccacc	268
QY	262	aaacacagactctcttagaagaagtacaagaagcactactcttgtaattctgtcaaccatgaatc	321
Db	2687	aaacacagactctcttagaagaagaacagaaagcactactctgtgaattctgtcaaccatgaatc	2744
QY	322	agagactctcttaaagaagcgcacgaatttgaaaggagaagatlaacatacttgggttaacgt	381
Db	2747	agagactctcttaaagaagcgcacgaatttgaaaggagaagatlaacatacttgggttaacgt	2800
QY	382	ttggggggggcgagcgatgctgtaaggaagcgcgaaagaanaactgtatccaagataccg	441
Db	2807	ttggggggggcgagcgatgctgtaaggaagcgcgaaagaanaactgtatccaagataccg	2866
QY	442	gccccaacccaagccaacaaagatgtagaagatttlttggaacagctgatttggcagactg	501
Db	2867	gccccaacccaagccaacaaagatgtagaagatttlttggaacagctgatttggcagactg	2920
QY	502	tgaatcccgagggtttgcgaccttagcagccccaacttaccggttaaccgaagaanaaaggg	561
Db	2927	tgaatcccgagggtttgcgaccttagcagccccaacttaccggttaaccgaagaanaaaggg	2986
QY	562	ggaattccctcgagctcctgagcaccgaagaagatattgatagtatacaaaaaggcctctgtg	621
Db	2987	ggaattccctcgagctcctgagcaccgaagaagatattgatagtatacaaaaaggcctctgtg	3048

QY 622 agcgcaactgctctgagccctccctagcgttaactaaaccccttaaccttattgtgtagag 681
DB 3047 agcgcaactgctctgagccctccctagcgttaactaaaccccttaaccttattgtgtagag 3106
QY 682 cgttaaggagtagcccgagagagttttaaccacaaaccccttagagccatgtgagagagctgtt 741
DB 3107 cgttaaggagtagcccgagagagttttaaccacaaaccccttagagccatgtgagagagctgtt 3166
QY 742 gcttaaccttcaaaagaaagcttgaatcctctgtagccagtggttggccgttagtctgaagct 801
DB 3167 gcttaaccttcaaaagaaagcttgaatcctctgtagccagtggttggccgttagtctgaagct 3226
QY 802 atcgagctgtgagccatctagcgtcaagagccgtgacaacatttgactcttgagacagaataa 861
DB 3227 atcgagctgtgagccatctagcgtcaagagccgtgacaacatttgactcttgagacagaataa 3286
QY 862 actgtaataagcccccaatgcatgtgagaaacatcgttcggcagccccagacccgatagtatg 921
DB 3287 actgtaataagcccccaatgcatgtgagaaacatcgttcggcagccccagacccgatagtatg 3346
QY 922 accaagcccgcatgaccccatatacaagctgtcttcaacagagaggttcaacttgcgt 981
DB 3347 accaagcccgcatgaccccatatacaagctgtcttcaacagagaggttcaacttgcgt 3406
QY 982 ccaacagccgcttcaacccctgcaactcttctgctgaagagactatgaacacagtgact 1041
DB 3407 ccaacagccgcttcaacccctgcaactcttctgctgaagagactatgaacacagtgact 3466
QY 1042 catgattgcatcaactattgattgagagagactggtgtccgcaagaccccttacagacata 1101
DB 3467 catgattgcatcaactattgattgagagagactggtgtccgcaagaccccttacagacata 3526
QY 1102 ccctgtactgagaaagtgtaacactgttcaactgagcaggaagcagctatgtggtgagaggt 1161
DB 3527 ccctgtactgagaaagtgtaacactgttcaactgagcaggaagcagctatgtggtgagaggt 3586
QY 1162 aagagagatgctggtggtggtgagtggtgagcagggagcccgacagatctggtgccaagcagctg 1221
DB 3587 aagagagatgctggtggtggtgagtggtgagcagggagcccgacagatctggtgccaagcagctg 3646
QY 1222 ccggaagaagacttcaagcgaagaagcgtgagctatgagccctcaacgaaggttgcggtg 1281
DB 3647 ccggaagaagacttcaagcgaagaagcgtgagctatgagccctcaacgaaggttgcggtg 3706
QY 1282 gccgaagaagacttcaataacattatatacagacagcagtgatgcttctgagactgacac 1341
DB 3707 gccgaagaagacttcaataacattatatacagacagcagtgatgcttctgagactgacac 3766
QY 1342 gtacacggtggtccatctatacaaaaggggtgtcttaacctcagcagggagggagaataaag 1401
DB 3767 gtacacggtggtccatctatacaaaaggggtgtcttaacctcagcagggagggagaataaag 3826
QY 1402 aacaaagaagaattcttaagccatttagaagccttaccatttgcacaaaagccttagttatt 1461
DB 3827 aacaaagaagaattcttaagccatttagaagccttaccatttgcacaaaagccttagttatt 3886
QY 1462 ataacagctcctgacatacgaagaagccaagaatctcatctagagaggaacccagatggtc 1521
DB 3887 ataacagctcctgacatacgaagaagccaagaatctcatctagagaggaacccagatggtc 3946
QY 1522 gaccgggttgcgaagcagcagccaggtgttaaccttctgcttaataatagaacgccc 1581
DB 3947 gaccgggttgcgaagcagcagccaggtgttaaccttctgcttaataatagaacacccc 4006
QY 1582 aaagccccagagacccagagacagtagaaccttagaagactgcaagaagataaaagaata 1641
DB 4007 aaagccccagagacccagagacagtagaaccttagaagactgcaagaagataaaagaata 4066
QY 1642 gaccagttctctgagactccggaagagactgtatatacctcatatgaggaaggaatcctg 1701
DB 4067 gaccagttctctgagactccggaagagactgtatatacctcatatgaggaaggaatcctg 4126

QY 1702 ccccaaaagaaggttagaaatctccaacagatatcgttcaacccactlagagact 1761
DB 4127 ccccaaaagaaggttagaaatctccaacagatatcgttcaacccactlagagact 4186
QY 1762 aaacactgacagcagttggttagaacaatcccttataatgttctgagagctacccaggtg 1821
DB 4187 aaacactgacagcagttggttagaacaatcccttataatgttctgagagctacccaggtg 4246
QY 1822 gctgactggtgtcaaacacttgtgtgtccctgacagctgtttaaigtcaatccttcaga 1881
DB 4247 gctgactggtgtcaaacacttgtgtgtccctgacagctgtttaaigtcaatccttcaga 4285
QY 1882 atacctcaaggaagaagactaaagggagccacccagcgctcactcgtggaggtgacttc 1941
DB 4286 ----- 4285
QY 1942 actgaggttaagccggttaatacgaagaacaatatcattgtgtttgttagacacttc 2001
DB 4286 -----ggtlaagccggttaatacgaagaacaatatcattgtgtttgttagacacttc 4340
QY 2002 tcagatggttagagcttactcactaagaagaagacttcaacgctgtgtgtagaaga 2061
DB 4341 tcagatggttagagcttactcactaagaagaagacttcaacgctgtgtgtagaaga 4400
QY 2062 atactgga-ggaatltttccaagatttggaatacctaaggttaataggtlcaagacatgtg 2120
DB 4401 atactggaaggaatltttccaagatttggaatacctaaggttaataggtlcaagacatgtg 4460
QY 2121 tccagcttctgttgcagaggttaagtcagagagctgccaagattatggaggtttagaa 2180
DB 4461 tccagcttctgttgcagaggttaagtcagagagctgccaagattatggaggtttagaa 4520
QY 2181 actgcatgtgcatcaagaccccaaggtcaggaagaaggttaagagaaatgaataacat 2240
DB 4521 actgcatgtgcatcaagaccccaaggtcaggaagaaggttaagagaaatgaataacat 4580
QY 2241 taaagaagcccttaccatattgacacagagagctgacatgattgattgaggttctcct 2300
DB 4581 taaagaagcccttaccatattgacacagagagctgacatgattgattgaggttctcctc 4640
QY 2301 gcccttctgcttctttaaaggttagaagaacaccccttggaacagtttggctgacccctatga 2360
DB 4641 gcccttctgcttctttaaaggttagaagaacaccccttggaacagtttggctgacccctatga 4700
QY 2361 attgctctacggggagaccccccgttggcagaatgtgcttgcacatagttgtatgt 2420
DB 4701 attgctctacggggagaccccccgttggcagaatgtgcttgcacatagttgtatgt 4760
QY 2421 gctgcttccagccttctgttcttagagctcaagagcgtctgaagtggtgagggcagc 2480
DB 4761 gctgcttccagccttctgttcttagagctcaagagcgtctgaagtggtgagggcagc 4820
QY 2481 gttggaagcagctccggagagccttactcctcaagctcaagagcgtctgaagtggtgagggcagc 2537
DB 4821 gttggaagcagctccggagagccttactcctcaagctcaagagcgtctgaagtggtgagggcagc 4880
QY 2538 ccaagttgagagattcaagttatgttagacgccaacggtgcaggaagaacccctcagagctggtg 2597
DB 4881 ccaagttgagagattcaagttatgttagacgccaacggtgcaggaagaacccctcagagctggtg 4940
QY 2598 gaagagacttatcttcgtaatttttagacacacacacggttggaaagtccgaagaataccc 2657
DB 4941 gaagagacttatcttcgtaatttttagacacacacacggttggaaagtccgaagaataccc 5000
QY 2658 caactgcatcattcatcccaagctttaaagccggcgcacactcccgattcgggttggaaagc 2717
DB 5001 caactgcatcattcatcccaagctttaaagccggcgcacactcccgattcgggttggaaagc 5060
QY 2718 cgaagaagctgaataatcccttaagcttcgctcactcagctgtgttcttactctgtcaa 2777
DB 5061 cgaagaagctgaataatcccttaagcttcgctcactcagctgtgttcttactctgtcaa 5120
QY 2778 taacttctaagtttaagttagaagccttgtgtgacagccggaactccataaaccttat 2837

```

Db      5121  taactcctcaagttaattgtaaacgctctgtgacacgcccgaactccataaaccttat 5180
Qy      2838  ctctacgtgttaacttaactactccgtgtacaggtattataataaacagcactaaagg 2897
Db      5181  ctctcaacttgtaacttaactgactccggtacaggtattataataaacagcactaaagg 5240
Qy      2898  aggtcccttggtggagcctgtgctgtgaattatataatgctgccttcgatacgaatcc 2957
Db      5241  aggtcccttggtggagcctgtgctgtgaattatataatgctgccttcgatacgaatcc 5300
Qy      2958  gtctcaatgacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3017
Db      5301  gtctcaatgacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 5360
Qy      3018  caggaccccccaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 3077
Db      5361  caggaccccccaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 5420
Qy      3078  ggaagtgtacataacttctaattgatgatgatgatgatgatgatgatgatgatgatgat 3137
Db      5421  ggaagtgtacataacttctaattgatgatgatgatgatgatgatgatgatgatgatgat 5480
Qy      3138  taagttaactcttttgttaacaactctacacgtataatcaatttaattatgcccattg 3197
Db      5481  taagttaactcttttgttaacaactctacacgtataatcaatttaattatgcccattg 5540
Qy      3198  gatgaaagattgtgcaacagcaggtacaaaagatgtacaaaatgaagaataagcgtgc 3257
Db      5541  gatgaaagattgtgcaacagcaggtacaaaagatgtacaaaatgaagaataagcgtgc 5600
Qy      3258  attcgttagacctagattacttaaaaaaattcattcaataaaaaaataaaaaaataaaa 3317
Db      5601  attcgttagacctagattacttaaaaaaattcattcaataaaaaaataaaaaaataaaa 5660

RESULT 5
AAV09699
ID      AAV09699 standard; DNA; 8196 BP.
XX
AC      AAV09699;
XX
DT      19-MAY-1998 (first entry)
XX
DE      Porcine retrovirus DNA encoding, GAG, POL and ENV.
XX
KW      Porcine retrovirus; PoEV; POL protein; ENV protein; GAG protein;
KM      vaccine; diagnosis; xenotransplantation; prophylactic; therapeutic; ds.
XX
OS      Porcine retrovirus.
XX
FH      Key
FT      CDS
FT      576..2126
FT      /tag= a
FT      /product= GAG protein
FT      /note= "Virion core polypeptide"
FT      2143..5733
FT      /tag= b
FT      /product= POL protein
FT      /note= "polymerase peptide sequence as given in
FT      specification"
FT      5606..7576
FT      /tag= c
FT      /product= ENV protein
FT      /note= "envelope protein"
XX
PN      MO9740167-A1.
XX
PD      30-OCT-1997.
XX
PF      18-APR-1997; 97WO-GB01087.
XX
PR      10-FEB-1997; 97GB-0002668.

```

```

PR      19-APR-1996; 96GB-0008164.
XX
PA      (IMUT-) IMUTRAN LTD.
PA      (OONE-) Q-OONE BIOTECH LTD.
PI      Galbraith DN, Haworth C, Lees GM, Smith KT;
XX
DR      WPI; 1997-535851/49.
XX
PT      Polynucleotide encoding porcine retrovirus expression product -
PT      useful to develop products for use in vaccines, diagnosis and
PT      xeno-transplantation
XX
PS      Claim 4; Fig 2; 69pp; English.
XX
CC      This DNA sequence encodes the porcine retrovirus (PoEV) virion core
CC      polypeptide (GAG), polymerase (POL) and envelope (ENV) proteins. These
CC      proteins can be used to develop viral vaccines, antisense nucleic acids,
CC      ribozymes and other antiviral agents. They can also be used in
CC      xeno-transplantation technology and as diagnostic tools.
XX
SQ      Sequence 8196 BP; 2165 A; 2061 C; 2147 G; 1820 T; 3 other;

```

```

Query Match      87.9%; Score 2916.8; DB 18; Length 8196;
Best local Similarity 96.0%; Pred. No. 0;
Matches 3013; Conservative 2; Mismatches 114; Indels 9; Gaps 2;

```

```

Qy      19  acagatgctctctctctgctgtgagatcacacccactagcaacacatttgccttcgaa 78
Db      2956  aaagatgctctctctctgctgtgagatcacacccactagcaacacatttgccttcgaa 3015
Qy      79  tgaagagatccaggtgtacggaagaacccggtgacgtctcacttgaaacccagctgccaaagg 138
Db      3016  tgaagagatccaggtgtacggaagaacccggtgacgtctcacttgaaacccagctgccaaagg 3075
Qy      139  ttcaagaactcccgaccactcttgcagaaagccctacacagagacctggccaactcagg 198
Db      3076  ttcaagaactcccgaccactcttgcagaaagccctacacagagacctggccaactcagg 3135
Qy      199  attcaaacacctcaagtgacacctctctccagtaagtgatgacgtctcttgcggagagcc 258
Db      3136  attcaaacacctcaagtgacacctctctccagtaagtgatgacgtctcttgcggagagcc 3195
Qy      259  accaaacagagactgtcttgaagaagtagcagaagcactacgtctggaattgtcgaactaagc 318
Db      3196  accaaacagagactgtcttgaagaagtagcagaagcactacgtctggaattgtcgaactaagc 3255
Qy      319  tacagagcctctgcttaagaagagccacagatttcagagaagagtaactacttgggtac 378
Db      3256  tacagagcctctgcttaagaagagccacagatttcagagaagagtaactacttgggtac 3315
Qy      379  agtttcggtggcggcgaagcagatgctgacaggaagcaggaagaanaactgtgtccagata 438
Db      3316  agtttcggtggcggcgaagcagatgctgacaggaagcaggaagaanaactgtgtccagata 4375
Qy      439  ccggcccccaacacagcaaaaagaatgagagatgttttgggagacgctggagttttggcga 498
Db      3376  ccggcccccaacacagcaaaaagaatgagagatgttttgggagacgctggagttttggcga 3435
Qy      499  ctgtgataccgggtgttgcgacacttgcacacccactctaccccgactaaccaagaanaaa 558
Db      3436  ctgtgataccgggtgttgcgacacttgcacacccactctaccccgactaaccaagaanaaa 3495
Qy      559  ggggattctctctgtggtctcgtgagcaccagaagcagatttgatgtatacaaaaaggcctg 618
Db      3496  ggggattctctctgtggtctcgtgagcaccagaagcagatttgatgtatacaaaaaggcctg 3555
Qy      619  ctgaagcagacctgctcgtgacctccctgaagtaactaaaccttaaccttatgtgat 678
Db      3556  ctgaagcagacctgctcgtgacctccctgaagtaactaaaccttaaccttatgtgat 3615
Qy      679  gagcgtlaaggagtagcccgagagattttaaccacaacaccttaggacatggagagagact 738

```

Db 3616 gagcgttaaggagtagcccgagagagtttaacccaacacccctagacacatggaagagacct 3675
QY 739 gttgccaactgtccaagaagcttgatctcgttagcagatggttggccgatctctgaag 798
Db 3676 gtgcgccaactgtccaagaagcttgatcccgtagccagtggttggccgatagtccgaag 3735
QY 799 gctatcgcaactgtgcgcatctgtccaagaagcgttgacaatcttgcttggaacgaat 858
Db 3736 gctatcgcaactgtgcgcatctgtccaagaagcgttgacaatcttgcttggaacgaat 3795
QY 859 ataacgttaataccccccatgtcatcttgaggaaacatcgcttcggagcccccagacgaatg 918
Db 3796 ataacgttaataccccccatgtcatcttgaggaaacatcgcttcggagcccccagacgaatg 3855
QY 919 atagaacacgcccgcatacgaacacacttcaaacgctgtcttcacagaagaggtcacttcc 978
Db 3856 atagaacacgcccgcatacgaacacacttcaaacgctgtcttcacagaagaggtcacttcc 3915
QY 979 gctccacacgcccgcatacgaacacacttctgcctgaagaagacatgaacacgaatg 1038
Db 3916 gctccacacgcccgcatacgaacacacttctgcctgaagaagacatgaacacgaatg 3975
QY 1039 actcatgatctgcatacactatgtatgaggaagactggtgctccgacgaagaccttaacagc 1098
Db 3976 actcatgatctgcatacactatgtatgaggaagactggtgctccgacgaagaccttaacagc 4035
QY 1099 ataccgctgaactggaagaagctgaactgctgaactgaacgaagcagcctatgctgagaa 1158
Db 4036 ataccgctgaactggaagaagctgaactgctgaactgaacgaagcagcctatgctgagaa 4095
QY 1159 ggttaagaagatgctgctggcgccagctggtgagacgggacccgacgaactctgggacagac 1218
Db 4096 ggttaagaagatgctgctggcgccagctggtgagacgggacccgacgaactctgggacagac 4155
QY 1219 ctgcgcgaagaagacttcaagcgaacaaagcgtgaactatgcccctcaacggaagcttgccg 1278
Db 4156 ctgcgcgaagaagacttcaagcgaacaaagcgtgaactatgcccctcaacggaagcttgccg 4215
QY 1279 ctgcgcgaagaagacttcaacataacatttatacgaacagcagctatgcttgcgactga 1338
Db 4216 ctgcgcgaagaagacttcaacataacatttatacgaacagcagctatgcttgcgactga 4275
QY 1339 caagctacaacgggcatctatacaaaagggtgtgttaactcagcagaaggaggaata 1398
Db 4276 caagctacaacgggcatctatacaaaagggtgtgttaactcagcagaaggaggaata 4335
QY 1399 aagacaagaagaaattctaaagcctattagaagccttaacatttgcacaaagaagctagct 1458
Db 4336 aagacaagaagaaattctaaagcctattagaagccttaacatttgcacaaagaagctagct 4395
QY 1459 attatacactgtctcgtgaacatcagaagaagcccaagatctatatactagaggaacacagatg 1518
Db 4396 attatacactgtctcgtgaacatcagaagaagcccaagatctatatactagaggaacacagatg 4455
QY 1519 gctgacacgggttgccaagcagcagcagcggcgttaaacctctgtgcttaataagaag 1578
Db 4456 gctgacacgggttgccaagcagcagcagcggcgttaaacctctgtgcttaataagaag 4515
QY 1579 ccccaagccccagaaacccagacagctacacacctaagaagactgtaagaagataaagaag 1638
Db 4516 ccccaagccccagaaacccagacagctacacacctaagaagactgtaagaagataaagaag 4575
QY 1639 atagacagctctctgagactccggaagggaacctgtatacctcatactggaaggaatc 1698
Db 4576 atagacagctctctgagactccggaagggaacctgtatacctcatactggaaggaatc 4635
QY 1699 ctgcccccaaaaagaaggttagaataatgtccaacagataacatcgtttaacccaactagga 1758
Db 4636 ctgcccccaaaaagaaggttagaataatgtccaacagataacatcgtttaacccaactagga 4695
QY 1759 actaaacacctgacagatgtgtcagaacatccctcatcatgttcttgagctaccaga 1818
Db 4696 actaaacacctgacagatgtgtcagaacatccctcatcatgttcttgagctaccaga 4755
QY 1819 gctgctgaactcgtgttcaaacatgtgtgtccctgcagcgtgtgttaagttaactcctcc 1878
Db 4756 gctgctgaactcgtgttcaaacatgtgtgtccctgcagcgtgtgttaagttaactcctcc 4815
QY 1879 agaatacctccaagaagaagactaagggaagccacccagcgtcactggaagctgagac 1938
Db 4816 agaatacctccaagaagaagactaagggaagccacccagcgtcactggaagctgagac 4875
QY 1939 ttcactgtgttaagcccggttaaaacgaacaaatctcattggtttttgtagaacc 1998
Db 4876 ttcactgtgttaagcccggttaaaacgaacaaatctcattggtttttgtagaacc 4955
QY 1999 ttctcagaatgtgtagaagcttatcctactaagaagaagacttcaacccgtgtgtgcaag 2058
Db 4936 ttctcagaatgtgtagaagcttatcctactaagaagaagacttcaacccgtgtgtgcaag 4995
QY 2059 aaaatactggaagaaattttcccaagatttgaatacctaaagtaaatggttcagaacat 2118
Db 4996 aaaatactggaagaaattttcccaagatttgaatacctaaagtaaatggttcagaacat 5055
QY 2119 ggtccagcttctgttgcgaagtaagtcaggaagactggtccagaatatttgaggtatgtg 2178
Db 5056 ggtccagcttctgttgcgaagtaagtcaggaagactggtccagaatatttgaggtatgtg 5115
QY 2179 aaactgcatctgtcatalacagaccccaagctcagaagcgtgtagaagagatgaatagaac 2238
Db 5116 aaactgcatctgtcatalacagaccccaagctcagaagcgtgtagaagagatgaatagaac 5175
QY 2239 attaaaggaaccccttaacaaatttgacccaagaagactgcatataatgtatgtgacctc 2298
Db 5176 attaaaggaaccccttaacaaatttgacccaagaagactgcatataatgtatgtgacctc 5255
QY 2299 ctgccccttggtctttagagtgtagaagaccccttgacagcttggagctaacccctat 2358
Db 5256 ctgccccttggtctttagagtgtagaagaccccttgacagcttggagctaacccctat 5295
QY 2359 gaattgctctacgggggaccccccccgcttggcagaatctgcttgcacatagctgat 2418
Db 5296 gaattgctctacgggggaccccccccgcttggcagaatctgcttgcacatagctgat 5355
QY 2419 gttgctgttcccgaccttctctctaggtctaaaggcgtcgggtgtgtagacaga 2478
Db 5356 gttgctgttcccgaccttctctctaggtctaaaggcgtcgggtgtgtagacaga 5415
QY 2479 gctgtgaagcagctccgggaagcctactcagaagagagactgtcacaatctccacatcgctc 2538
Db 5416 gctgtgaagcagctccgggaagcctactcagaagagagactgtcacaatctccacatcgctc 5475
QY 2539 caagtgtgaagatcagctatgtttagacgcacggctgcaggaacacccogagaactcggtg 2598
Db 5476 caagtgtgaagatcagctatgtttagacgcacggctgcaggaacacccogagaactcggtg 5555
QY 2599 aagggaaccttactcgtacttctgaaccaacacacggctgtggaagctcgaaggaatcccc 2658
Db 5536 aagggaaccttactcgtacttctgaaccaacacacggctgtggaagctcgaaggaatcccc 5595
QY 2659 acctgtgatalcgtacatcccaagcttaagcggcgccacactcccgatctcggtgtggaagcc 2718
Db 5596 acctgtgatalcgtacatcccaagcttaagcggcgccacactcccgatctcggtgtggaagcc 5555
QY 2719 gaagaagctgaataatcccttaagctcgtccatcagctggtgttcttactcgtcaat 2778
Db 5656 gaagaagctgaataatcccttaagctcgtccatcagctggtgttcttactcgtcaat 5715
QY 2779 aactcctaagttaatgttaaacgctgtgagaccccggaactccataaaccttatc 2838
Db 5716 aactcctaagttaatgttaaacgctgtgagaccccggaactccataaaccttatc 5775
QY 2839 tctcacctggttacttaagctccggtacagctatataatataatcaacacactcaagga 2898
Db 5776 tctcacctggttacttaagctccggtacagctatataatataatcaacacactcaagga 5835

CC This DNA sequence encodes the porcine retrovirus (POEV) virion core
CC polypeptide (GAG), polymerase (POL) and envelope (ENV) proteins and
CC also includes the Long Terminal Repeat (LTR). These proteins can be used
CC to develop viral vaccines, antisense nucleic acids, ribozymes and other
CC antiviral agents. They can also be used in xeno-transplantation
CC technology and as diagnostic tools.

XX Sequence 8209 BP; 2168 A; 2064 C; 2154 G; 1823 T; 0 other:

Query Match 87.8%; Score 2916; DB 18; Length 8209;
Best Local Similarity 96.0%; Pred. No. 0;
Matches 3014; Conservative 0; Mismatches 115; Indels 9; Gaps 2;

QY 19 acagatgctctctctgctgagattacacccactagccaacacttttgccttcga 78
DB 2970 aaagatgctctctctgctgagattacacccactagccaacacttttgccttcga 3029
QY 79 tggagagatccaggtacggaagaaacggcgacgtacactggaaccgactgcccgaagg 138
DB 3030 tggagagatccaggtacggaagaaacggcgacgtacactggaaccgactgcccgaagg 3089
QY 139 ttcaagaactcccgacacatcttgaagaaacccctacacagggactggccaactcagg 198
DB 3090 ttcaagaactcccgacacatcttgaagaaacccctacacagggactggccaactcagg 3149
QY 199 atccaacacctaggtgacacctctccagtagctgtagactgtcttcggggagacc 258
DB 3150 atccaacacctaggtgacacctctccagtagctgtagactgtcttcggggagacc 3209
QY 259 accaacaagagctgcttaagaaggtacgaagcactactgtggaattgtctgacctaagc 318
DB 3210 accaacaagagctgcttaagaaggtacgaagcactactgtggaattgtctgacctaagc 3269
QY 319 tacagagctctctgtaagaagcccaagattgcaagagagaggtacatacttgggtac 378
DB 3270 tacagagctctctgtaagaagcccaagattgcaagagagaggtacatacttgggtac 3329
QY 379 agtttgggggcgagcagcagtagctgtaaggaagcagcaggaagaaactgtagtccagata 438
DB 3330 agtttgggggcgagcagcagtagctgtaaggaagcagcaggaagaaactgtagtccagata 3389
QY 439 ccggcccccaacacagcacaacaaagtgagagagtttttggggacagcttgatttgcaga 498
DB 3390 ccggcccccaacacagcacaacaaagtgagagagtttttggggacagcttgatttgcaga 3449
QY 499 ctgtggaatcccggtgttggagacttagcagccccactctaacccgttaaccaaagaaaa 558
DB 3450 ctgtggaatcccggtgttggagacttagcagccccactctaacccgttaaccaaagaaaa 3509
QY 559 ggggagattccctgagctctgagcacaagaagcatcttatgctatcaaaagacccctg 618
DB 3510 ggggagattccctgagctctgagcacaagaagcatcttatgctatcaaaagacccctg 3569
QY 619 ctgagcgacactgctctgcccctccctgacgttaactaaaccccttaccctttatgtgat 678
DB 3570 ctgagcgacactgctctgcccctccctgacgttaactaaaccccttaccctttatgtgat 3629
QY 679 gagcgttaaggaatgagccgaggaaggtttaaccacaacccctaggaacatggaagaaacct 738
DB 3630 gagcgttaaggaatgagccgaggaaggtttaaccacaacccctaggaacatggaagaaacct 3689
QY 739 gttccctaactgttcaaaagagctgtgactcgttgaagcagtggttggccgtatgtctgaag 798
DB 3690 gttccctaactgttcaaaagagctgtgactcgttgaagcagtggttggccgtatgtctgaag 3749
QY 799 gctctcagctgttggccatactgttcaagaaagcgtgacaacattgactttgggaacgaat 858
DB 3750 gctctcagctgttggccatactgttcaagaaagcgtgacaacattgactttgggaacgaat 3809
QY 859 ataacttaataagcccccaatgcatgtgagaacatcgttggcgagcccccaagccgaatg 918
DB 3810 ataacttaataagcccccaatgcatgtgagaacatcgttggcgagcccccaagccgaatg 3869

QY 919 atgaccaagcccgcatgaccactatcaaaagcctgtcttcacagagaggtcacttcc 978
DB 3870 atgaccaagcccgcatgaccactatcaaaagcctgtcttcacagagaggtcacttcc 3929
QY 979 gttccaccagccgcttcaaccctgcaactctctctgctgtaagaagactgatagaacagtg 1038
DB 3930 gttccaccagccgcttcaaccctgcaactctctctgctgtaagaagactgatagaacagtg 3989
QY 1039 actatgttgcatacactatgtatgagaaagctgggtgcgcaagagacttaccagac 1098
DB 3990 actatgttgcatacactatgtatgagaaagctgggtgcgcaagagacttaccagac 4049
QY 1099 ataccgctgactggaagagtgtaaacctggtctacatgacccgcaagcagctatgtgtgaa 1158
DB 4050 ataccgctgactggaagagtgtaaacctggtctacatgacccgcaagcagctatgtgtgaa 4109
QY 1159 ggtaaagagatgcttggtggcgagtggtggaacggacccgacagatctggccaagac 1218
DB 4110 ggtaaagagatgcttggtggcgagtggtggaacggacccgacagatctggccaagac 4169
QY 1219 ctggcggaaagaaacttcaagcgcaaaagcgtgagctcagtcagccacgcaagcttgcgg 1278
DB 4170 ctggcggaaagaaacttcaagcgcaaaagcgtgagctcagtcagccacgcaagcttgcgg 4229
QY 1279 ctggcggaaagaaatccataaacaattatacgaacagcaggtatgaccttgcactgca 1338
DB 4230 ctggcggaaagaaatccataaacaattatacgaacagcaggtatgaccttgcactgca 4289
QY 1339 caagtacacggggccatctataaanaaggggtgtgtcttacttaccgaaggaaggaata 1398
DB 4290 caagtacacggggccatctataaanaaggggtgtgtcttacttaccgaaggaaggaata 4349
QY 1399 aagaacaagaggaattctaaagcctatataaagccttacttggccaanaagctagct 1458
DB 4350 aagaacaagaggaattctaaagcctatataaagccttacttggccaanaagctagct 4409
QY 1459 attatacactgtctctgacatcagaagaacccaagaatctcalatcagaaggaacccaagt 1518
DB 4410 attatacactgtctctgacatcagaagaacccaagaatctcalatcagaaggaacccaagt 4469
QY 1519 gctgacccgggttgcgaagcagcagccagcgtgttaaccttctgacctataatagaacg 1578
DB 4470 gctgacccgggttgcgaagcagcagccagcgtgttaaccttctgacctataatagaacg 4529
QY 1579 cccaagcccccaagacccaagacagctacacccctagaagaactggccaagagataaaaaag 1638
DB 4530 cccaagcccccaagacccaagacagctacacccctagaagaactggccaagagataaaaaag 4589
QY 1639 atagaacagttctctgagactcccggaaggagacctgtactactcatatgggaaggaatc 1698
DB 4590 atagaacagttctctgagactcccggaaggagacctgtactactcatatgggaaggaatc 4649
QY 1699 ctggcccaaaagaagaggttgaataatgttccaagagatatacgtctcaacccaactaga 1758
DB 4650 ctggcccaaaagaagaggttgaataatgttccaagagatatacgtctcaacccaactaga 4709
QY 1759 actaaacacttgcagcagttgtgtcagaacatcccttcatatgttctgagcttaccagga 1818
DB 4710 actaaacacttgcagcagttgtgtcagaacatcccttcatatgttctgagcttaccagga 4769
QY 1819 gtggtgactcgtgtgtcaaacattgtgtgcccctgcagcgtgtgtaatgacttacccttc 1878
DB 4770 gtggtgactcgtgtgtcaaacattgtgtgcccctgcagcgtgtgtaatgacttacccttc 4829
QY 1879 agaatactccaggaagaagactaaagggaagcccccagggcgtcactgtggaatgtgac 1938
DB 4830 agaatactccaggaagaagactaaagggaagcccccagggcgtcactgtggaatgtgac 4889
QY 1939 ttcaactgagtaagccggtctaaataacgaaacaaatctatgtgttcttctgtagaac 1998
DB 4890 ttcaactgagtaagccggtctaaataacgaaacaaatctatgtgttcttctgtagaac 4949

the histological localisation of activated retrovirus. Using Polymerase Chain Reaction DNA Quantitation (PDQ) on blood mononuclear cells, infectivity titration and susceptibility testing can be performed. Ultimately animal donors without intact porcine retroviral sequences or a lower copy number of viral elements could be selected.

Sequence 7892 BP; 2191 A; 1915 C; 1980 G; 1806 T; 0 other;

Query Match 87.8%; Score 2915.2; DB 18; Length 7892;
Best Local Similarity 94.3%; Pred. No. 0;
Matches 3051; Conservative 0; Mismatches 173; Indels 12; Gaps 2;

QY 22 gatgacctctctcgtgagattacacccaccagcaaacacttttgcctcgaatg 81
DB 2970 gatgacctctctcgtgagattacacccaccagcaaacacttttgcctcgaatg 3029
QY 82 agagatccaggtacggaagaaacgggagctacacttgacccgactgcccgaagggtc 141
DB 3030 agagatccaggtacggaagaaacgggagctacacttgacccgactgcccgaagggtc 3089
QY 142 aagaacctcccgacatctttgacgaagccctacacagggacttgccaattcaagatc 201
DB 3090 aagaacctcccgacatctttgacgaagccctacacagggacttgccaattcaagatc 3149
QY 202 caaacacctcaggtgacccctccctccagtaagtgatgacactgctctgagcgagacacc 261
DB 3150 caaacacctcaggtgacccctccctccagtaagtgatgacactgctctgagcgagacacc 3209
QY 262 aaacagagactgtctagaaggtacgaaggaacactcgtcgtgaattgtctgacctagctac 321
DB 3210 aaacagagactgtctagaaggtacgaaggaacactcgtcgtgaattgtctgacctagctac 3269
QY 322 agagacctctctcgtgagagggcccaagatttgaggaagagagtaacatcttgagggtacagt 381
DB 3270 agagacctctctcgtgagagggcccaagatttgaggaagagagtaacatcttgagggtacagt 3329
QY 382 ttgagcgagcgagcgacgtatgctgacgagacgacggaagaaactgtaagtcagacacg 441
DB 3330 ttgagcgagcgagcgacgtatgctgacgagacgacggaagaaactgtaagtcagacacg 3389
QY 442 gccccaacacacagccaacaaacgaagagagtttttggggacagctgattttcagacgt 501
DB 3390 gccccaacacacagccaacaaacgaagagagtttttggggacagctgattttcagacgt 3449
QY 502 tgagatcccgaggtttgacacctagacgcccactacacgccttaacacaaagaaagag 561
DB 3450 tgagatcccgaggtttgacacctagacgcccactacacgccttaacacaaagaaagag 3509
QY 562 ggaattctctgagctcctgagacacgaagagcaattgatactcaaaaagggcctgctg 621
DB 3510 ggaattctctgagctcctgagacacgaagagcaattgatactcaaaaagggcctgctg 3569
QY 622 agcgacacctgctctgagccctccctgagcttaactaaacctttaccctttatgtggatgag 681
DB 3570 agcgacacctgctctgagccctccctgagcttaactaaacctttaccctttatgtggatgag 3629
QY 682 cgtlaagagagtagccgagagagtttlaacccaacacccctagagacatgagaaagacgtgtc 741
DB 3630 cgtlaagagagtagccgagagagtttlaacccaacacccctagagacatgagaaagacgtgtc 3689
QY 742 gctactactgtcaaaagagctgtgactcgttagccagtggttggccgtatgtctgaaggtc 801
DB 3690 gctactactgtcaaaagagctgtgactcgttagccagtggttggccgtatgtctgaaggtc 3749
QY 802 atcgcaactgtgagccatagctgtaagaagacgctgacaaatgtactttggagacgaatata 861
DB 3750 atcgcaactgtgagccatagctgtaagaagacgctgacaaatgtactttggagacgaatata 3809
QY 862 acttaataagcccccatgcatctgagagaacatcgttcgacagcccccaagccagatgagt 921
DB 3810 acttaataagcccccatgcatctgagagaacatcgttcgacagcccccaagccagatgagt 3869

QY 922 accaagcccgcatgaccacatataaagcctgctctccacagaagaggtcattcgtc 981
DB 3870 accaagcccgcatgaccacatataaagcctgctctccacagaagaggtcattcgtc 3929
QY 982 ccacagccgctctcaacctgccaactctctgctcgtgaagagactgtatgaaccggtact 1041
DB 3930 ccacagccgctctcaacctgccaactctctgctcgtgaagagactgtatgaaccggtact 3989
QY 1042 catgattccatcaactttgatitgagagactgggtgcgcgaagagaccttaacagata 1101
DB 3990 catgattccatcaactttgatitgagagactgggtgcgcgaagagaccttaacagata 4049
QY 1102 ccgctgacttggaagagtgctaacctggttcaactcgtacggaagcagctatgtgtggaaggt 1161
DB 4050 ccgctgacttggaagagtgctaacctggttcaactcgtacggaagcagctatgtgtggaaggt 4109
QY 1162 aagaagatggtcgtggcgagctgtgtggaacggcgacagatcttgggcagacgctg 1221
DB 4110 aagaagatggtcgtggcgagctgtgtggaacggcgacagatcttgggcagacgctg 4169
QY 1222 ccggaaggaacttcagcgcaaaagcttgagctcagtcagccctcaagacttgcgctg 1281
DB 4170 ccggaaggaacttcagcgcaaaagcttgagctcagtcagccctcaagacttgcgctg 4229
QY 1282 gccgaagggaaatccataaacttatacgggaacagagtgatgcctttgacgtgcacac 1341
DB 4230 gccgaagggaaatccataaacttatacgggaacagagtgatgcctttgacgtgcacac 4289
QY 1342 gtacacgggggccaatataaacaaggggtgtgtcttacttaacgaggaaggaataaag 1401
DB 4290 gtacacgggggccaatataaacaaggggtgtgtcttacttaacgaggaaggaataaag 4349
QY 1402 aacaagaggaatcttaagccatataaagccttatacttgccaaaaaggtactatt 1461
DB 4350 aacaagaggaatcttaagccatataaagccttatacttgccaaaaaggtactatt 4409
QY 1462 atacactgtctctgagacatccagaagacaaagatctcatatctagaagggaacacagtgct 1521
DB 4410 atacactgtctctgagacatccagaagacaaagatctcatatctagaagggaacacagtgct 4469
QY 1522 gacccggtttgccaagcagcagccagagctgttaacctcttcgacctatataagaaacgcc 1581
DB 4470 gacccggtttgccaagcagcagccagagctgttaacctcttcgacctatataagaaacgcc 4529
QY 1582 aaagccccaagaaacccaagacagatcacacctagaagacgtggcaagagataaaagata 1641
DB 4530 aaagccccaagaaacccaagacagatcacacctagaagacgtggcaagagataaaagata 4589
QY 1642 gaccaggttctctgagactccgggaaggagcgtctatcctcaatgtgggaagaaatcctg 1701
DB 4590 gaccaggttctctgagactccgggaaggagcgtctatcctcaatgtgggaagaaatcctg 4649
QY 1702 ccccaaaaagaaaggtttagaataatgtccaaacagatacatctgctlaacccacttagaact 1761
DB 4650 ccccaaaaagaaaggtttagaataatgtccaaacagatacatctgctlaacccacttagaact 4709
QY 1762 aaacacctgacagagtggttgaagaacatccctcatcatgtctcgaaggtaccagaggt 1821
DB 4710 aaacacctgacagagtggttgaagaacatccctcatcatgtctcgaaggtaccagaggt 4769
QY 1822 gctgactcgtgtgttaaaatgtgtgaccttcagagcgtgttaagtcataatccctcaga 1881
DB 4770 gctgactcgtgtgttaaaatgtgtgaccttcagagcgtgttaagtcataatccctcaga 4829
QY 1882 atacctccagaaagagactaagggaagccacacagcgctcactctggaagtgagcttc 1941
DB 4830 atacctccagaaagagactaagggaagccacacagcgctcactctggaagtgagcttc 4889
QY 1942 actgaggtaaagccggcttaaatcaacggaacaaatcatatgtgttttgtgacacctt 2001
DB 4890 actgaggtaaagccggcttaaatcaacggaacaaatcatatgtgttttgtgacacctt 4949
QY 2002 tcaggaatggttgaaggtcttactctactaagaagagacttaacccgtgtgtgctagaaga 2061

```

Db 4950 tcaagatggttagaggtctactactactaagaagaagacttcaaccggttggtcaaaaaa 5009
Qy 2062 atacggggaattttccaagatttggaatbactaaggaatgaaggtcagacaatggt 2121
Db 5010 atactggagaagaattttccaagatttggaatbactaaggaatgaaggtcagacaatggt 5069
Qy 2122 ccaagcttcgttgcacaggaagtaagtaaggaaggtcagacaatbactaaggaatggtgaa 2181
Db 5070 ccaagcttcgttgcacaggaagtaagtaaggaaggtcagacaatbactaaggaatggtgaa 5129
Qy 2182 ctgcatctgtcatatacagaaccccaagctcagagacaggttaagagagatgaataacatt 2241
Db 5130 ctgcatctgtcatatacagaaccccaagctcagagacaggttaagagagatgaataacatt 5189
Qy 2242 aaagagaccccttacaatctagacaagagactgagcatatgatttgatgagctctctg 2301
Db 5190 aaagagaccccttacaatctagacaagagactgagcatatgatttgatgagctctctg 5249
Qy 2302 ccccttgcgtctttagggtaggaacaaccccttgacaagtttgagctgagccctatgaa 2361
Db 5250 ccccttgcgtctttagggtaggaacaaccccttgagacagtttgagctgagccctatgaa 5309
Qy 2362 ttgctctacgggggagaccccccgttggcagaatttgcttgcacatagtctgcatgtg 2421
Db 5310 ttactctacgggggagacccccccttgtagaattgctctgtacatagtctgcatgtg 5369
Qy 2422 ctgcttccacagccttgcgtcttgcgtacagcagcgtcagagtggtggagcgacgagcg 2481
Db 5370 ctgcttccacagccttgcgtcttgcgtacagcagcgtcagagtggtggagcgacgagcg 5429
Qy 2482 tgaagacagctcgcggagcgctactc--aggaagagactgtgacagttccacatcgttc 2538
Db 5430 tgaagacagctcgcggagcgctactcagagagagagagactgtgacagttccacatcgttc 5489
Qy 2539 caagtctgagagattcaagtctatgtttagagcgccacggtgcaggaacccctgcagctg 2598
Db 5490 caagtctgagagattcaagtctatgtttagagcgccacggtgcaggaacccctgcagctg 5549
Qy 2599 aaggagactatctcgtactctttagacacacacacagcggttgaagtcggaagaatcccc 2658
Db 5550 aaggagactatctcgtactctttagacacacacacagcggttgaagtcggaagaatcccc 5609
Qy 2659 acctgagatcagatccacacggttaagcgcgcacccctgcagttcggggttgaaagcc 2718
Db 5610 acctgagatcagatccacacggttaagcgcgcacccctgcagttcggggttgaaagcc 5669
Qy 2719 gaaaagactgaataatcccccttaagctcgcctcagtcaggttgcttccctactgtcaat 2778
Db 5670 gaaaagactgaataatcccccttaagctcgcctcagtcaggttgcttccctactgtcaat 5729
Qy 2779 aacctcctcaagtttaagtgaacgcgttgcgacagccgaactccatcaataaacctatc 2838
Db 5730 aacctcctcaagtttaagtgaacgcgttgcgacagccgaactccatcaataaacctatc 5789
Qy 2839 tctcaacctgttacttactactcgcgttacaggtlatatataacagcacaacgaaggga 2898
Db 5790 tctcaacctgttacttactactcgcgttacaggtlatatataacagcacaacgaaggga 5849
Qy 2899 ggcctccttgggagacctggtgcgcctgaatatatgtctgcctcagtcagttacccctg 2958
Db 5850 ggcctccttgggagacctggtgcgcctgaatatatgtctgcctcagtcagttacccctg 5909
Qy 2959 tctcaatgacacagcgcacaccccgatgacgcgtgcttgaaggttgaaggttggccc 3018
Db 5910 tct-----gacctcaccccccagatactcccatgtccaagggatttcttggccc 5960
Qy 3019 aggaaccccaataatgaagaatattgtggaatctccagagatttcttgcagaagcag 3078
Db 5961 aggaaccccaataatgaagaatattgtggaatctccagagatttcttgcagaagcag 6020
Qy 3079 gagctgataacttctatgatgtggaattggaaatgcagttctccagagaagcagagt 3138

```

```

Db 6021 gaactgttaaccttaatgatgatattggaatgccaacctccagcagataggt 6080
Qy 3139 aagttactcttctgttaacaatccctacacgttatcaatcaattatgagcagtgag 3198
Db 6081 aagttcttctgttaacaacccatacagctctgagcaattatcaactgacccgagat 6140
Qy 3199 atgaaagatttgcacacagcggttacaaaagatgtaacgaataaagaatgagct 3254
Db 6141 tagaactggaagcccaagtgctctctctcagacactagattacttaataaataagt 6196

RESULT 8
AAF77727
ID AAF77727 standard; DNA; 8132 BP.
XX
AC AAF77727;
XX
DE 23-MAY-2001 (first entry)
XX
DE Nucleotide sequence of a retrovirus found in miniature swine.
XX
KW Retrovirus; graft transplantation; xenotransplantation; miniature swine;
KW ss.
XX
XX Unidentified.
XX
XX US6190861-B1.
XX
XX 20-FEB-2001.
XX
XX 13-DEC-1996; 96US-0766528.
XX
XX 14-DEC-1995; 95US-0572645.
XX
XX (GENO ) GEN HOSPITAL CORP.
XX
XX Fishman JA;
XX
XX WPI: 2001-256211/26.
XX
XX P-PSDB: AAB73285, AAB73286, AAB73287.
XX
XX Assessing risk of endogenous retroviruses in clinical practice and in
XX xenotransplantation, comprises using probe sequences derived from swine
XX or miniature swine retroviral genome -
XX
XX Claim 1; Fig 3; 127pp; English.
XX
XX The present invention relates to a method for screening a cell or tissue
XX for the presence or expression of a retrovirus (RV), comprising
XX contacting a target nucleic acid from the cell or tissue with a second
XX nucleic acid from the present invention (e.g. the present sequence or a
XX fragment thereof). The method is useful for RV detection and to assess
XX graft transplantation risk. Screening of animals allows the elimination
XX of donors with active replication of known viruses. Inactive proviruses
XX can be detected and inactivated, allowing identification and elimination
XX of potential human pathogens derived from swine in a manner not possible
XX in the outbred human organ donor population and is important to the
XX development of human xenotransplantation.
XX
XX Sequence 8132 BP; 2248 A; 1977 C; 2037 G; 1870 T; 0 other;

Query Match 87.8%; Score 2915.2; DB 22; Length 8132;
Best Local Similarity 94.3%; Pred. No. 0;
Matches 3051; Conservative 0; Mismatches 173; Indels 12; Gaps 2;

Qy 22 gatgccttctctgctgagattacacccctcagcaacacactttgcttgcagatgg 81
Db 2970 gatgccttctctgctgagattacacccctcagcaacacactttgcttgcagatgg 3029
Qy 82 agagatcaggtacgggaagacgcggcagctacactggaacccgagctgcacaaaggttc 141
Db 3030 agagatcaggtacgggaagacgcggcagctacactggaacccgagctgcacaaaggttc 3089

```

QY 142 aagaactccccacatctttagaagaagccctacagaaggacttggccaacttcaagatc 201
|||||
Db 3090 aagaactccccacatctttagaagaagccctacagaaggacttggccaacttcaagatc 3149
QY 202 caaacaccttaagttgacccctccctccagttacgtatgtaaccttcttcttgcggaggaacc 261
|||||
Db 3150 caaacaccttaagttgacccctccctccagttacgtatgtaaccttcttcttgcggaggaacc 3209
QY 262 aaaaagagactgcttagaaggtagaaggacactactgtctgtaattgtcttaacctaagctac 321
|||||
Db 3210 aaaaagagactgcttagaaggtagaaggacactactgtctgtaattgtcttaacctaagctac 3269
QY 322 agagcccttctgtaagaagggccagatcttgagagagaggtaaacatacttgggttaccatg 381
|||||
Db 3270 agagcccttctgtaagaagggccagatcttgagagagaggtaaacatacttgggttaccatg 3329
QY 382 ttgcgggagcgagcagatgagctgagcagagacaggaagaaactgtatgtaacatacag 441
|||||
Db 3330 ttgcgggagcgagcagatgagctgagcagagacaggaagaaactgtatgtaacatacag 3389
QY 442 gcccccaacacagccaacaagaagtgagagagtttttgaggacagctgattttgacagctg 501
|||||
Db 3390 gcccccaacacagccaacaagaagtgagagagtttttgaggacagctgattttgacagctg 3449
QY 502 tggatcccgagggttttgacacttagcagcccaactacccgcttaacccaagaagaagg 561
|||||
Db 3450 tggatcccgagggttttgacacttagcagcccaactacccgcttaacccaagaagaagg 3509
QY 562 ggaattccctggggtctcttgagacccagaagacatttgatgtatcaaaaaggccctgctg 621
|||||
Db 3510 ggaattccctggggtctcttgagacccagaagacatttgatgtatcaaaaaggccctgctg 3569
QY 622 agcgaacctgctctgagccctcccttgagcttaacccaacttacccttlatgtgagatgag 681
|||||
Db 3570 agcgaacctgctctgagccctcccttgagcttaacccaacttacccttlatgtgagatgag 3629
QY 682 cgttaaggagtagcccgagagattttaacccaaccttagagacatgagagagactgtt 741
|||||
Db 3630 cgttaaggagtagcccgagagattttaacccaaccttagagacatgagagagactgtt 3689
QY 742 ggcctactgttcaagaagactttagcctgttagccagtggttgcgcgatalgtctgaaaggt 801
|||||
Db 3690 ggcctactgttcaagaagactttagcctgttagccagtggttgcgcgatalgtctgaaaggt 3749
QY 802 atcgcaagctgttgcacatactgttcaagaagcgttgcacaacttgaacttgggacagaatata 861
|||||
Db 3750 atcgcaagctgttgcacatactgttcaagaagcgttgcacaacttgaacttgggacagaatata 3809
QY 862 actgttaatagcccccaactgtatgtgagaacaatcgttccgagcccccaagaccgaatgag 921
|||||
Db 3810 actgttaatagcccccaactgtatgtgagaacaatcgttccgagcccccaagaccgaatgag 3869
QY 922 accaagcccgcatagaccacatacaagcctgtcttccacagagaggtcacttgcgt 981
|||||
Db 3870 accaagcccgcatagaccacatacaagcctgtcttccacagagaggtcacttgcgt 3929
QY 982 ccacacagcgctctcaacacctgccaactcttctgcttgaagaagactgatagaacagttgact 1041
|||||
Db 3930 ccacacagcgctctcaacacctgccaactcttctgcttgaagaagactgatagaacagttgact 3989
QY 1042 catgattgcatcactatattgattgagagagactgggttccgcaagaaccttaacagacata 1101
|||||
Db 3990 catgattgcatcactatattgattgagagagactgggttccgcaagaaccttaacagacata 4049
QY 1102 ccgcttgactgagagagtgtaaaccttggcttactgaagcagagcagctatgtgtggagaggt 1161
|||||
Db 4050 ccgcttgactgagagagtgtaaaccttggcttactgaagcagagcagctatgtgtggagaggt 4109
QY 1162 aagaagatgagctgggagcagtggttgaacgagacccgacagatcttgggacacagcctg 1221
|||||
Db 4110 aagaagatgagctgggagcagtggttgaacgagacccgacagatcttgggacacagcctg 4169

QY 1222 ccggaaggaaacttcagcgcgaagaagcttgagctcatalgtgcccacagcgaagcttgcgctg 1281
|||||
Db 4170 ccggaaggaaacttcagcgcgaagaagcttgcagctcatalgtgcccacagcgaagcttgcgctg 4229
QY 1282 gccgaaggaaatccataaacttatacgagcagcaggtatgtccttggagcgcagac 1341
|||||
Db 4230 gccgaaggaaatccataaacttatacgagcagcaggtatgtccttggagcgcagac 4289
QY 1342 gtacaggggacatctataaacaagaagggttgccttactcaagcagggaggaataaag 1401
|||||
Db 4290 gtacatggggacatctataaacaagaagggttgccttactcaagcagggaggaataaag 4349
QY 1402 aaaaaagagaaattcctaagcctatagaagcctatagattggccaaaaggctgactt 1461
|||||
Db 4350 aaaaaagagaaattcctaagcctatagaagcctatagattggccaaaaggctgactt 4409
QY 1462 atacactgctcctggagcatcagaagaacccaagaatcctatctatagaagggaacagatgct 1521
|||||
Db 4410 atacactgctcctggagcatcagaagaatcctatctatagaagggaacagatgct 4469
QY 1522 gacgggttgcgaagcagcagccagcctgtttaaacttctgctatataagaacgcgc 1581
|||||
Db 4470 gacgggttgcgaagcagcagccagcctgtttaaacttctgctatataagaatgccc 4529
QY 1582 aaagccccaagaacccaagacagatgacacccctagaagactggcgaagagataaagaata 1641
|||||
Db 4530 aaagccccaagaacccaagacagatgacacccctagaagactggcgaagagataaagaata 4589
QY 1642 gacaggttctctgagactcccgagggagcctgtctatcactcaatgaggaagaatacctg 1701
|||||
Db 4590 gacaggttctctgagactcccgagggagcctgtctatcactcaatgaggaagaatacctg 4649
QY 1702 ccccaagaagaaggtltaagaatgtccaacagatatacgtctaaacccactagaact 1761
|||||
Db 4650 ccccaagaagaaggtltaagaatgtccaacagatatacgtctaaacccactagaact 4709
QY 1762 aaaaactgctgacagattggttcaagaacatcccttcaatgcttcaaggctcacaagggtg 1821
|||||
Db 4710 aaaaactgctgacagattggttcaagaacatcccttcaatgcttcaaggctcacaagggtg 4769
QY 1822 gctgactcgttgttcaaaacatgtgttgccttgcagcctggttaatgtcaatccctcaga 1881
|||||
Db 4770 gctgactcgttgttcaaaacatgtgttgccttgcagcctggttaatgtcaatccctcaga 4829
QY 1882 atacctccaagaagaagactaaagggaagccaacccagcgtcactggtgaagttgacttc 1941
|||||
Db 4830 atacctccaagaagaagactaaagggaagccaacccagcgtcactggtgaagttgacttc 4889
QY 1942 acttgagtgtaagccggcttaaatagcgaagaacaatatctattggttcttggtagaaccttt 2001
|||||
Db 4890 acttgagtgtaagccggcttaaatagcgaagaacaatatctattggttcttggtagaaccttt 4949
QY 2002 tcaagatggtagaggtctatctcactaagaagaagacttcaacccgttggctagaaga 2061
|||||
Db 4950 tcaagatggtagaggtctatctcactaagaagaagacttcaacccgttggctagaaga 5009
QY 2062 atacttgaggaatcttcccaagatttgaataacttaaggttaataaggttcagacaatgct 2121
|||||
Db 5010 atacttgaggaatcttcccaagatttgaataacttaaggttaataaggttcagacaatgct 5069
QY 2122 ccaagcttctgttgcgaagtgtaagtcagagactggccaagatacttgggattgattggaaa 2181
|||||
Db 5070 ccaagcttctgttgcgaagtgtaagtcagagactggccaagatacttgggattgattggaaa 5129
QY 2182 ctgcatctgtgcatatagaaccacaagctcagagacaggttagagaggaatgaatagaacatc 2241
|||||
Db 5130 ctgcatctgtgcatatagaaccacaagctcagagacaggttagagaggaatgaatagaacatc 5189
QY 2242 aaagaaaccttacaataattgacacagagactggcatataatgatttgatgctcctg 2301
|||||
Db 5190 aaagaaaccttacaataattgacacagagactggcatataatgatttgatgctcctg 5249
QY 2302 cccttgtgctttaggttgaaggtaaaccccccttgagacaggttgggtctgacccccatgaa 2361

Db	5250	cccttctgcttcttaaggttagaacaacccctcgacagtttggtgccccctatgaa	5309
Oy	2362	ttgcctctacggggagaccccccgcttgycagaatatgtgcctttgcacatagtgtcgtatgtg	2421
Db	5310	ttactctacggggagacccccccatctgtglaagaattgtctctgcatacatagttcgtacgtg	5369
Oy	2422	ctgcttcccgagcctttgttctctaaagctcaagcgctcgagttggttgtagcgaagacg	2481
Db	5370	ctgcttcccgagcctttgttctctaaagctcaagcgctcgagttggttgtagaacaacgagcg	5429
Oy	2482	tggaaagcagctccggagagcctactc---agggagagacttgcgaattccacatcgcttc	2538
Db	5430	tggagagcaactccggagagcgctactcaagcgagggagacttgcagatcccaatcgtttc	5489
Oy	2539	caagtttggaagattcagtctctatgtttaagcgcacccgcttgcaggaacacctcgagactcgtgtg	2598
Db	5490	caagttggagagattcaagctctacgtcttaagcgcacccgctgcaggaacacctcgagactcgtgtg	5549
Oy	2599	aaggagaccttctcgcgaactttgcacacacacacacggcgcttgaaagctcgaaagaaatccc	2658
Db	5550	aaggagccctatctcgcgaactttgcacacacacacacggcgcttgaaagctcgaaagaaatccc	5609
Oy	2659	acctgtgatccatgatcccaagcttaagccggcgccacacctcccgatctcggggttggaagctc	2718
Db	5610	acctgtgatccatgatcccaagcttaagccggcgccacacctcccgatctcggggttggaagctc	5669
Oy	2719	gaaaagagctgaaatcccccttaagcttcgcgcctccatcgcggtgtgtcttactctgtcaat	2778
Db	5670	gaaaagagctgaaatcccccttaagcttcgcgcctccatcgcggtgtgtcttactctgtcaat	5729
Oy	2779	aaccctccaggttaatgtgtaaacgctcttggaacgccccgaactcccaataacaccttatc	2838
Db	5730	aacctccaggttaatgtgtaaacgctcttggaacgccccgaactcccaataacaccttatc	5789
Oy	2839	tctcaacctgttactactgaactcggtacagaglatataataltaaacagactccaagggga	2898
Db	5790	tctcaacctgttactactgaactcggtacagaglatataataltaaacagactccaagggga	5849
Oy	2899	ggctcccttgggagccggcggtgcgtgaattatatagtcttcgcttcgatagaatccctcgg	2958
Db	5850	ggctcccttgggagccggcggtgcgtgaattatatagtcttcgcttcgatagaatccctcgg	5909
Oy	2959	tctcaatlgacagagcacacaccccccgatgatctcgttactcgttgtaagggctttaaagcttgcgc	3018
Db	5910	tct-----gacctacacccccagatatcctcgaatgcgaagatttatgtttgcgc	5960
Oy	3019	aggaagcccccaataatagaagaatatgtgvgaaatccccaagatctcctttgcaagcaatg	3078
Db	5961	aggaagcccccaataatagaagaatatgtgvgaaatccccaagatctcctttgcaagcaatg	6020
Oy	3079	gagctgagataactctcaatgatgaggaattgggaatggcgccagctcttcgaagcaagcaagt	3138
Db	6021	gagctgagataactctcaatgatgaggaattgggaatggcgccagctcttcgaagcaagcaagt	6080
Oy	3139	aagttactctttgttaacaatctaccagttataatcaatlttaattatggtcagagag	3198
Db	6081	aagttactctttgttaacaacaactataccagctctgacaaatttaattacctgacctgat	6140
Oy	3199	atggaagaagtatggcaacgacgggtgtaacaagaatgtatgaagaataagcaataaagct	3254
Db	6141	tagaacttggaaagcccaagatgctctcctctcgaaccttgaacttatacctaataaataaatt	6196

RESULT	ID	SEQ	ID NO:
9	AAC67022	standard; DNM; 4918 bp.	22
XX	AAC67022;		
XX	27-MAR-2001 (first entry)		
XX	PERV env protein coding sequence		

XX	Xenotransplantation; infectious agent; vaccine; ds.
KW	
OS	Porcine endogenous retrovirus.
XX	
PN	WO200071726-A1.
XX	
PD	30-NOV-2000.
XX	
PF	24-MAY-2000; 2000MO-US14296.
XX	
PR	24-MAY-1999; 99US-0135631.
XX	
PA	(MAYO-) MAYO MEDICAL VENTURES.
XX	
PI	Federspiel MJ;
XX	
DR	WPI; 2001-032041/04.
XX	
PT	Inhibiting or preventing infectious agent transmission in mammalian
PT	transplant recipients, by introducing recombinant DNA comprising DNA
PT	encoding extracellular proteins of the agent into donor cells, such as
PT	swine cells -
XX	
ES	
XX	Claim 16; Page 109-111; 144pp; English.
XX	
CC	The present invention provides a method to prevent the transmission of
CC	infectious agents during xenotransplantation. This involves introducing
CC	to donor swine cells a recombinant DNA encoding a peptide fragment from
CC	the infectious agent, and then introducing these cells into the
CC	transplant recipient.
XX	
XX	
SO	Sequence 4918 BP; 1382 A; 1173 C; 1183 G; 1180 T; 0 other;

Query Match	82.0%	Score 2721.4	DB 22	Length 4918
Best Local Similarity	97.7%	Pred. No. 0		
Matches 2771	Conservative	0	Mismatches 61	Indels 3
				Gaps 1
QY 486	tgagatttcacagactgtgatcccgagggttttcgaccttagacgagccacctacacgct	545		
Db 1	tgagatttcgagactgtgatcccgagggttttcgaccttagacgagccacctacacgct	60		
QY 546	aaccaaaagaaaaaggggagattctcctgggtccctgaagcaaccagaagcatltagatcat	605		
Db 61	aaccaaaagaaaaaggggagattctcctgggtccctgaagcaaccagaagcatltagatcat	120		
QY 606	caaaaagggcctgtcgaaagcagccacgcgtctggccctccctgaagcaactaaacctttac	665		
Db 121	caaaaagggcctgtcgaaagcagccacgcgtctggccctccctgaagcaactaaacctttac	180		
QY 666	ccttattgtgtagcgtaagggatagcccgaggaatttaaccccaaccttagagacc	725		
Db 181	ccttattgtgtagcgcgtaagggatagcccgaggaatttaaccccaaccttagagacc	240		
QY 726	atgagagagaccttttgctacctgtccaagaagttagatcctgtagccagtggttgacc	785		
Db 241	atgagagagagaccttttgctacctgtccaagaagttagatcctgtagccagtggttgacc	300		
QY 786	cgatagtcgaagcgatcagctgtgagccatctgtgtcaaggaagcgctgtgcaaatgtac	845		
Db 301	cgatagcctgaagcgatcagctgtgagccatctgtgtcaaggaagcgctgtgcaaatgtac	360		
QY 846	tttggagcagaatataactgttaatatagccccccatgcatltagagaacatcgttcggcagcc	905		
Db 361	tttggagcagaatataactgttaatatagccccccatgcatltagagaacatcgttcggcagcc	420		
QY 906	cccgacacggaatggaatgacaaacgcccgcgaatgacccaatataaaagcctgtcttccaaaga	965		
Db 421	cccgacacggaatggaatgacaaacgcccgcgaatgacccaatataaaagcctgtcttccaaaga	480		
QY 966	gagagtaacttgcgtccaccagccggtctcaacccctgcacatctctgtcctgaagagac	1025		

Db 481 gagggtcaagctctgcctcacacgacgcgtctcaacccctgcacactctctgcctgaagagac 540
QY 1026 tgaatgaacccagtgatctatgaattgcctcaactatttgatggagagctgggtccgcga 1085
Db 541 tgaatgaacccagtgatctatgaattgcctcaactatttgatggagagctgggtccgcga 600
QY 1086 ggaacttacaacatacagctgcgtgagtgagtgatctaaacttggtctcaactgaagaaacag 1145
Db 601 ggaacttacaacatacagctgcgtgagtgagtgatctaaacttggtctcaactgaagaaacag 660
QY 1146 ctatctgtgtgaagtaagagatgctgcgtggcgacgtggtgagacggacccgcagat 1205
Db 661 ctatctgtgtgaagtaagagatgctgcgtggcgacgtggtgagacggacccgcagat 720
QY 1206 ctgggccaagcagctgcgtggagaaacttgaacggcaaaaggtctgaagctcagccctcac 1265
Db 721 ctgggccaagcagctgcgtggagaaacttgaacggcaaaaggtctgaagctcagccctcac 780
QY 1266 gcaagcttgctgcgtgcgcgaaggaatccataacattatatacagacagagtaagc 1325
Db 781 gcaagcttgctgcgtgcgcgaaggaatccataacattatatacagacagagtaagc 840
QY 1326 ctctgcagactgcacacgtaacacggcgccatctataacaaagggtgtgcttaccctcagc 1385
Db 841 ctctgcagactgcacacgtaacacggcgccatctataacaaagggtgtgcttaccctcagc 900
QY 1386 agggagggaaataaagaaacaaaggaaattctgaagcttataaagccttacccttgcgc 1445
Db 901 agggagggaaataaagaaacaaaggaaattctgaagcttataaagccttacccttgcgc 960
QY 1446 aaaaagcctagctatactatacactgctcctgcgacatcacagaaagccaaagatcatatag 1505
Db 961 aaaaagcctagctatactatacactgctcctgcgacatcacagaaagccaaagatcatatag 1020
QY 1506 agggaaacagatgctgcgcgggtgtgcgaacagagccagccagctgttacccttgcgc 1565
Db 1021 agggaaacagatgctgcgcgggtgtgcgaacagagccagccagctgttacccttgcgc 1080
QY 1566 tataatagaacacgccccaaagccccagaaacccagacagatcaccccttaagaagactgcga 1625
Db 1081 tatgatagaacacacgccccaaagccccagaaacccagacagatcaccccttaagaagactgcga 1140
QY 1626 agagaataaaaagaatagacagctctctgcgactccggagggagcctgctatacctcata 1685
Db 1141 agagaataaaaagaatagacagctctctgcgactccggagggagcctgctatacctcata 1200
QY 1686 tgggaaaggaatcctgcgcgcacaaagaaaggttagaataatgtccaaacagatacgtct 1745
Db 1201 tgggaaaggaatcctgcgcgcacaaagaaaggttagaataatgtccaaacagatacgtct 1260
QY 1746 aaccacactaggaacttaaacacctgcacagatgctgcagaaactcccttatacgttct 1805
Db 1261 aaccacactaggaacttaaacacctgcacagatgctgcagaaactcccttatacgttct 1320
QY 1806 gaggctaccagagatgctgcgtcgtgtgtaacaacattgtgtgcctgcagctggttaa 1865
Db 1321 gaggctaccagagatgctgcgtcgtgtgtaacaacattgtgtgcctgcagctggttaa 1380
QY 1866 tgcctatccttccaaataactccagagaaagagactaaggagagccacccagcgctca 1925
Db 1381 tgcctatccttccaaataactccagagaaagagactaaggagagccacccagcgctca 1440
QY 1926 ctgggaagtgaactcactgcagtgtaaacggcgtataatacagaaacaaatactatgt 1985
Db 1441 ctgggaagtgaactcactgcagtgtaaacggcgtataatacagaaacaaatactatgt 1500
QY 1986 ttctgtagaacaccttccaagatgggtagaagcttatacctcctaagaagagacttcaac 2045
Db 1501 ttctgtagaacaccttccaagatgggtagaagcttatacctcctaagaagagacttcaac 1560
QY 2046 gctgtgtgctaaagaaatactgcaggaatcttccaagatttgaaatacctaaggtat 2105
Db 1561 gctgtgtgctaaagaaatactgcaggaatcttccaagatttgaaatacctaaggtat 1620

QY 2106 agggtcagaaatggttcacagcttctgctgcccaggtaagtcaaggagctgcgccaagata 2165
Db 1621 agggtcagaaatggttcacagcttctgctgcccaggtaagtcaaggagctgcgccaagata 1680
QY 2166 ggggattgattggaaactgcattgtgcatacagaccccaagctcaagacagtgtagag 2225
Db 1681 ggggattgattggaaactgcattgtgcatacagaccccaagctcaagacagtgtagag 1740
QY 2226 gatgaataaaccatataaagagagaccccttaccatattgaccacagagactggtcataagc 2285
Db 1741 gatgaataaaccatataaagagagaccccttaccatattgaccacagagactggtcataagc 1800
QY 2286 ttgagtgctctcctgcgccttgcgtctttaggggtgaggaacacccctgcagcttgg 2345
Db 1801 ttgagtgctctcctgcgccttgcgtctttaggggtgaggaacacccctgcagcttgg 1860
QY 2346 gctgacccctatgaattgctctcagggagaccccccccgcttggcagaaatgctcttgc 2405
Db 1861 gctgacccctatgaattgctctcagggagaccccccccgcttggcagaaatgctcttgc 1920
QY 2406 acatagtgctgattgctctcttccagccttctctcctaggtcaagcgctcgatg 2465
Db 1921 acatagtgctgattgctctcttccagccttctctcctaggtcaagcgctcgatg 1980
QY 2466 ggtgagcgagcgagcggtggaagcagctccggagagcctacac--aggaagagacttgca 2522
Db 1981 ggtgagcgagcgagcggtggaagcagctccggagagcctacagagagagagacttgca 2040
QY 2523 agtccacatcgtcttccaaagttagaatactgataatgtaaaagccaaagcggtgga 2582
Db 2041 gatacccaactcgttccaaagttagaatactgataatgtaaaagccaaagcggtgga 2100
QY 2583 cctcgagactcgtgtggaagggagactatctcgtacttcttgacacacaaagcggtgga 2642
Db 2101 cctcgagactcgtgtggaagggagactatctcgtacttcttgacacacaaagcggtgga 2160
QY 2643 agtcgaaggaatccccaactgatacagatcacatcccaagcgttgaagccacccctccga 2702
Db 2161 agtcgaaggaatccccaactgatacagatcacatcccaagcgttgaagccacccctccga 2220
QY 2703 ttctgggggtggaagggagaaagacgtgaataatcccttaagcttgcctccatcggtgt 2762
Db 2221 ttctgggggtggaagggagaaagacgtgaataatcccttaagcttgcctccatcggtgt 2280
QY 2763 tccctactcgttaataactcctccaagttaatggtlaaacgcttgcgaagcccgact 2822
Db 2281 tccctactcgttaataactcctccaagttaatggtlaaacgcttgcgaagcccgact 2340
QY 2823 cccataaaccttatctctcaactcgtgttacttactgaactcccggtlaaggtataatatta 2882
Db 2241 cccataaaccttatctctcaactcgtgttacttactgaactcccggtlaaggtataatatta 2400
QY 2883 acagacataagggggagcctccttgggagccgtggtgcctgataatgctgcctc 2942
Db 2401 acagacataagggggagcctccttgggagccgtggtgcctgataatgctgcctc 2460
QY 2943 gatcagtaatccctgtgtctcaatgacagagccacaccccgatgtaactcgtgttagc 3002
Db 2461 gatcagtaatccctgtgtgtcaatgacagagccacaccccgatgtaactcgtgttagc 2520
QY 3003 ggttttaagcttgcgcagagacccccaataatgaagaataatgtggaatccctcagatt 3062
Db 2521 ggttttaagcttgcgcagagacccccaataatgaagaataatgtggaatccctcagatt 2580
QY 3063 tcccttgcaggaatgagctgcatataactcttaagtgaatggaatggaatgagct 3122
Db 2581 tcccttgcaggaatgagctgcatataactcttaagtgaatggaatggaatgagct 2640
QY 3123 ctcaagcaagacagatgaatgtaactcttgtttaaacaatccctacagttataatcaattta 3182
Db 2641 ctcaagcaagacagatgaatgtaactcttgtttaaacaatccctacagttataatcaattta 2700

QY 3183 attatgacatgagatggaagaattgcaacacgaggtacaaaagatgtaccgaata 3242
 |||||
 Db 2701 attatgacatgagatggaagaattgcaacacgaggtacaaaagatgtaccgaata 2760
 |||||
 QY 3243 agcaataagctgtcattcgttagaacttaacttaaaaataagttcactaaanaa 3302
 |||||
 Db 2761 agcaataagctgtcattcgttagaacttaacttaaaaataagttcactaaanaa 2820
 |||||
 QY 3303 aaaaaaanaanaa 3317
 |||||
 Db 2821 gaaacaagaanaa 2835
 |||||
 RESULT 10
 AAT74811
 ID AAT74811 standard; CDNA: 8060 BP.
 XX
 AC AAT74811;
 XX
 DT 11-FEB-1998 (first entry)
 XX
 DE Porcine retrovirus Tsukuba-1 CDNA.
 XX
 KW Retrovirus; porcine; GAG protein; POL protein; ENV protein;
 KW xenotransplantation; infectious; provirus; organ transplant; donor;
 KW activated virus; Tsukuba-1; PCR; ss.
 XX
 OS Porcine retrovirus.
 XX
 FH Key Location/Qualifiers
 FT CDS 86..2002
 FT /*tag= a
 FT /note= "putative ENV protein"
 FT 312..4686
 FT /*tag= b
 FT /note= "putative GAG protein"
 FT 4871..8060
 FT /*tag= c
 FT /note= "putative POL protein (partial)"
 XX
 PN W09721836-A1.
 XX
 PD 19-JUN-1997.
 XX
 PE 13-DEC-1996; 96WO-US19680.
 XX
 PR 14-DEC-1995; 95US-0572645.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PI Fishman JA;
 XX
 DR WPI; 1997-332804/30.
 XX
 PT New nucleic acid from porcine retroviruses - used for detecting
 PT viruses in transplant or other tissue and for assessing risk of
 PT transmitting infection to graft recipient
 XX
 PS Claim 1; Fig 1; 128pp; English.
 XX
 CC This sequence represents the purified porcine retroviral cDNA
 CC sequence of Tsukuba-1 and contains the putative coding regions for viral
 CC proteins GAG, POL and ENV. This sequence and PCR fragments generated
 CC from the sequence (see AAT74812-r74882) could be used to screen organs
 CC for porcine retroviruses prior to xenotransplantation. Transplantation
 CC can increase the likelihood of retroviral activation if intact and
 CC infectious proviruses are present. The porcine retroviral sequence can be
 CC used to generate probes to determine the level (e.g. copy number) of
 CC intact (i.e. potentially replicating) porcine provirus sequences in a
 CC strain of xenograft transplantation donors. It can be used to detect
 CC mutations, genetic lesions or viral recombinants and to determine the
 CC histological localisation of activated retrovirus. Using Polymerase Chain
 CC Reaction DNA Quantitation (PQ) on blood mononuclear cells, infectivity

CC titration and susceptibility testing can be performed. Ultimately animal
 CC donors without intact porcine retroviral sequences or with a lower copy
 CC number of viral elements could be selected.

XX Sequence 8060 BP; 2233 A; 1959 C; 2012 G; 1856 T; 0 other;

Query Match 72.9%; Score 2420.6; DB 18; Length 8060;
 Best Local Similarity 97.5%; Pred. No. 0;
 Matches 2502; Conservative 0; Mismatches 59; Indels 6; Gaps 4;

QY 22 gatgctctctcgtcgtgagattacacccacacagcagcaccatttctccttcagtg 81
 |||||
 Db 5495 gatgctctctcgtcgtgagattacacccacacagcagcaccatttctccttcagtg 5554
 |||||
 QY 82 agagatccaggtacggaagaacccgagcagctcactctgacccagctgcccagggctc 141
 |||||
 Db 5555 agagatccaggtacggaagaacccgagcagctcactctgacccagctgcccagggctc 5614
 |||||
 QY 142 aagaactcccgagcactctttagcagaagccctacacagagacccctgccaattcagatc 201
 |||||
 Db 5615 aagaactcccgagcactctttagcagaagccctacacagagacccctgccaattcagatc 5674
 |||||
 QY 202 caacacccctcaggtacccctcctccagtagcgtgagctcgtctgaggagcacc 261
 |||||
 Db 5675 caacacccctcaggtacccctcctccagtagcgtgagctcgtctgaggagcacc 5734
 |||||
 QY 262 aaacaggaactgcttagaaggtacgaagcactactctggaattgtctgactaaggtac 321
 |||||
 Db 5735 aaacaggaactgcttagaaggtacgaagcactactctggaattgtctgactaaggtac 5794
 |||||
 QY 322 agagcctctgcttagaagaagccagatttcaggaagaggttaacacttgggtacagt 381
 |||||
 Db 5795 agagcctctgcttagaagaagccagatttcaggaagaggttaacacttgggtacagt 5854
 |||||
 QY 382 ttgcgggggagcagcagctgctgacgagcagcagcagcagcagcagcagcagcagc 441
 |||||
 Db 5855 ttgcgggggagcagcagctgctgacgagcagcagcagcagcagcagcagcagcagc 5914
 |||||
 QY 442 gccccaacacacagcacaacaaagttagagaacttctgggagcagctggaatttgcagact 501
 |||||
 Db 5915 gccccaacacacagcacaacaaagttagagaacttctgggagcagctggaatttgcagact 5974
 |||||
 QY 502 tggatccggggttttgcagcacttagagccacacttacccgtacccaanaaaggagg 561
 |||||
 Db 5975 tggatccggggttttgcagcacttagagccacacttacccgtacccaanaaaggagg 6034
 |||||
 QY 562 ggaattccttgggtcctcctagacacaggaagcatttgaatctatacaaaaagccctgct 621
 |||||
 Db 6035 ggaattccttgggtcctcctagacacaggaagcatttgaatctatacaaaaagccctgct 6094
 |||||
 QY 622 agcgcacctgctcctgagccctccctgaactaactaaccttacccttattgtgagtag 681
 |||||
 Db 6095 agcgcacctgctcctgagccctccctgaactaactaaccttacccttattgtgagtag 6154
 |||||
 QY 682 cctaaaggagtagcccggaaggttttaaccacaccccttaggacatgagggagcacttt 741
 |||||
 Db 6155 cctaaaggagtagcccggaaggttttaaccacaccccttaggacatgagggagcactttc 6214
 |||||
 QY 742 gctactcttcaaaagaagcttgaatccctgtagcaggttggccgctatgtctgaagct 801
 |||||
 Db 6215 gctactcttcaaaagaagcttgaatccctgtagcaggttggccgctatgtctgaagct 6274
 |||||
 QY 802 atcgcagctctggtcactactgttcaagaagcgttgaacaaatgacttggagc-agaatat 860
 |||||
 Db 6275 atcgcagctctggtcactactgttcaagaagcgttgaacaaatgacttggagc-agaatat 6334
 |||||
 QY 861 aacttaagcccccatgacttggagaacttgggaagcccccagccagatgat 920
 |||||
 Db 6335 aacttaagcccccatgacttggagaacttgggaagcccccagccagatgat 6394
 |||||
 QY 921 gaccacgcccgcatagacacactcaaaagctgtcttcacagagaggtcactttgcg 980
 |||||

CC for the presence or expression of a retrovirus (RV), comprising
 CC contacting a target nucleic acid from the cell or tissue with a second
 CC nucleic acid from the present invention (e.g., the present sequence or a
 CC fragment thereof). The method is useful for RV detection and to assess
 CC graft transplantation risk. Screening of animals allows the elimination
 CC of donors with active replication of known viruses. Inactive proviruses
 CC can be detected and inactivated, allowing identification and elimination
 CC of potential human pathogens derived from swine in a manner not possible
 CC in the outbred human organ donor population and is important to the
 CC development of human xenotransplantation.

XX Sequence 8060 BP; 2233 A; 1959 C; 2012 G; 1856 T; 0 other;

Query Match 72.9%; Score 2420.6; DB 22; Length 8060;
 Best Local Similarity 97.5%; Pred. No. 0;
 Matches 2502; Conservative 0; Mismatches 59; Indels 6; Gaps 4;

QY 22 gatgctctctctgctgagatlaacacccactagccacacatttctgctcgaatg 81
 Db |||||||
 Db 5495 gatgctctctctgctgagatlaacacccactagccacacatttctgctcgaatg 5554
 QY 82 agaatatcaggttagcggagaagacccggcagctcaactccgagccagatcccaagggttc 141
 Db |||||||
 Db 5555 agaatatcaggttagcggagaagacccggcagctcaactccgagccagatcccaagggttc 5614
 QY 142 aagaatcccccagacatctttagcagagccctacacagagacctggtcccaattcagatc 201
 Db |||||||
 Db 5615 aagaatcccccagacatctttagcagagccctacacagagacctggtcccaattcagatc 5674
 QY 202 caacacccctcagctgacccctcctcagctgagatgagctgctctcgcggagagccacc 261
 Db |||||||
 Db 5675 caacacccctcagctgacccctcctcagctgagatgagctgctctcgcggagagccacc 5734
 QY 262 aaacagagacgttagaaggtacgaagagcactctgctggaattgtctgacaggttac 321
 Db |||||||
 Db 5735 aaacagagacgttagaaggtacgaagagcactctgctggaattgtctgacaggttac 5794
 QY 322 agagcctctgttaagaaagccagatttgcagagagaggtlaacatactcgggggtacat 381
 Db |||||||
 Db 5795 agagcctctgttaagaaagccagatttgcagagagaggtlaacatactcgggggtacat 5854
 QY 382 ttggggggcggcgacgtgctgacgagagcagcgaagaaactgtatgccagataccg 441
 Db |||||||
 Db 5855 ttggggggcggcgacgtgctgacgagagcagcgaagaaactgtatgccagataccg 5914
 QY 442 gcccaacccacagccaacaagtgagagagtttttggggacagcttgatttgcagactg 501
 Db |||||||
 Db 5915 gcccaacccacagccaacaagtgagagagtttttggggacagcttgatttgcagactg 5974
 QY 502 tggatcccggtgttgacacttagcagcccaactacccggttaaccgaagaagaagg 561
 Db |||||||
 Db 5975 tggatcccggtgttgacacttagcagcccaactacccggttaaccgaagaagaagg 6034
 QY 562 ggaattcctcgtgctctgaacacagaaagcaattgattgctatcaaaaagccctgctg 621
 Db |||||||
 Db 6035 ggaattcctcgtgctctgaacacagaaagcaattgattgctatcaaaaagccctgctg 6094
 QY 622 agcgcacctgctctgcccctccttgacgttaactaaagcccttaaccccttatgtggtatg 681
 Db |||||||
 Db 6095 agcgcacctgctctgcccctccttgacgttaactaaagcccttaaccccttatgtggtatg 6154
 QY 682 cgttaaggaatagcccgagaggttttaaccgaacccctagacacatgagagagactgct 741
 Db |||||||
 Db 6155 cgttaaggaatagcccgagaggttttaaccgaacccctagacacatgagagagactgct 6214
 QY 742 gacctactgtcacaagaagcttgatctgttagcagtggttggccgtatgtctgaagct 801
 Db |||||||
 Db 6215 gacctactgtcacaagaagcttgatctgttagcagtggttggccgtatgtctgaagct 6274
 QY 802 atcgcaactgtgacatactgtcgaagagcgtgacaaattgactttggac-agaatt 860
 Db |||||||
 Db 6275 atcgcaactgtgacatactgtcgaagagcgtgacaaattgactttggacagaataat 6334

QY 861 aactgtaaatagcccccatgcatcttgagagaacatcgttgcgagcccccagaccgatgat 920
 Db |||||||
 Db 6335 aactgtaaatagcccccatgcatcttgagagaacatcgttgcgagcccccagaccgatgat 6394
 QY 921 gaccacagcccgatagaccacatatacaagcctgctctcacaagagaggtcactctgc 980
 Db |||||||
 Db 6395 gaccacagcccgatagaccacatatacaagcctgctctcacaagagaggtcactctgc 6454
 QY 981 tccacacagccgctcctaaccctgacactcttctgcttgaagagactgatagaaccagtac 1040
 Db |||||||
 Db 6455 tccacacagccgctcctaaccctgacactcttctgcttgaagagactgatagaaccagtac 6514
 QY 1041 tcatgattgcacatcaactttagttagagagctgggtccgcaagagacccttacagacat 1100
 Db |||||||
 Db 6515 tcatgattgcacatcaactttagttagagagctgggtccgcaagagacccttacagacat 6574
 QY 1101 accgctgactggagaagtgctaaactgcttcaactgacggagaagcagatgtgtgaag 1160
 Db |||||||
 Db 6575 accgctgactggagaagtgctaaactgcttcaactgacggagaagcagatgtgtgtgaag 6634
 QY 1161 taagagatggtctggggcgagtggtgagacgggacccgacgactctggccagacgct 1220
 Db |||||||
 Db 6635 taagagatggtctggggcgagtggtgagacgggacccgacgactctggccagacgct 6694
 QY 1221 gccggagagaaacttcaagcgcaaaagcctgagctcaatggccctcacgcaagcttgcgct 1280
 Db |||||||
 Db 6695 gccggagagaaacttcaagcgcaaaagcctgagctcaatggccctcacgcaagcttgcgct 6754
 QY 1281 ggcggaagggaatccataaacaattatagcagacagagtagtgccttgcagctgacaca 1340
 Db |||||||
 Db 6755 ggcggaagggaatccataaacaattatagcagacagagtagtgccttgcagctgacaca 6814
 QY 1341 cgtacacggggccatctataaacaagggtgtgttaccttaagcagaaggaggaataaa 1400
 Db |||||||
 Db 6815 cgtacacggggccatctataaacaagggtgtgttaccttaagcagaaggaggaataaa 6874
 QY 1401 gaacaaagggagaatcttcaagcctatlaagaaccttlaacttgcgaanaaaggctgcat 1460
 Db |||||||
 Db 6875 gaacaaagggagaatcttcaagcctatlaagaaccttlaacttgcgaanaaaggctgcat 6934
 QY 1461 tatacactgtccttgatcatcagaagccaagatcatalactagaagggaacagatgac 1520
 Db |||||||
 Db 6935 tatacactgtccttgatcatcagaagccaagatcatalactagaagggaacagatgac 6994
 QY 1521 tgaccgggttgcgaagcggcgacggccaggtgttaacttctgctatataagaacgc 1580
 Db |||||||
 Db 6995 tgaccgggttgcgaagcggcgacggccaggtgttaacttctgctatataagaacgc 7054
 QY 1581 caaagcccccagaacccaagacgacgtacacccctagaagaactgccaagagataaagaat 1640
 Db |||||||
 Db 7055 caaagcccccagaacccaagacgacgtacacccctagaagaactgccaagagataaagaat 7114
 QY 1641 agacagatctcttgagactcggagggagcctgtcatcctcatatgtggaaaggaatcct 1700
 Db |||||||
 Db 7115 agacagatctcttgagactcggagggagcctgtcatcctcatatgtggaaaggaatcct 7173
 QY 1701 gccccaacaagaaggttagaataatgtccaac-agaatacactgcttaaccacactaggaa 1759
 Db |||||||
 Db 7117 gccccaacaagaaggttagaataatgtccaac-agaatacactgcttaaccacactaggaa 1723
 QY 1714 gccccaacaagaaggttagaataatgtccacaagaatatacatcgtctaaaccacactaggaa 7233
 Db |||||||
 Db 1716 gccccaacaagaaggttagaataatgtccacaagaatatacatcgtctaaaccacactaggaa 1819
 Db |||||||
 Db 7234 gccccaacaagaaggttagaataatgtccacaagaatatacatcgtctaaaccacactaggaa 7293
 QY 1820 tggctgactcgggtgctaaacattgtgtgcctcggcagctggttaattgataccttcca 1879
 Db |||||||
 Db 7294 tggctgactcgggtgctaaacattgtgtgcctcggcagctggttaattgataccttcca 7353
 QY 1880 gaatacctccaggaagagactaaagggaagccacccagggcgtcactgggaagtggact 1939
 Db |||||||
 Db 7354 gaatacctccaggaagagactaaagggaagccacccagggcgtcactgggaagtggact 7413

QY	1940	tcactcgtggtgtaaaagccgcgctcaaaatacggaaacaaatcatctatgtttttttagaagcact	1999
Db	7414	tcactcgtggtgtaaaagccgcgctcaaaatacggaaacaaatcatctatgtttttttagaagcact	7473
QY	2000	tttcaggaatggtgtagaggtctactctactcaagaagaagacttcaaccgtgtgtgtcaaga	2059
Db	7474	tttcaggaatggtgtagaggtctactctactcaagaagaagacttcaaccgtgtgtgtcaaga	7533
QY	2060	aaatactcggaggaatttttccaaagatttggaaattcgaatgaagttaagtgtcagaacaatg	2119
Db	7534	aaatactcggaggaatttttccaaagatttggaaattcgaatgaagttaagtgtcagaacaatg	7593
QY	2120	gtccagacttcctgtgtgcccaagttagacgaagactgtgcgaagaatctatgtgggaattgtatga	2179
Db	7594	gtccagacttcctgtgtgtgcccaagttagacgaagactgtgcgaagaatctatgtgggaattgtatga	7653
QY	2180	aactgcgatgtgtcatatagacacccccaaagctcagagacaggttagagagatgaatagaacca	2239
Db	7654	aactgcgatgtgtcatatagacacccccaaagctcagagacaggttagagagatgaatagaacca	7713
QY	2240	ttaaagagacccttaccacaatttaccacagagactgtgcattatattgtatgtgtctcc	2299
Db	7714	ttaaagagacccttaccacaatttaccacagagactgtgcattatattgtatgtgtgtctcc	7773
QY	2300	tgcaccttgcgtcttttaagggtgaggaacaccccttgcagacttggctgtgacccctcatg	2359
Db	7774	tgcaccttgcgtcttttaagggtgaggaacaccccttgcagacttggctgtgacccctcatg	7833
QY	2360	aattgcctctacggggagcccccccgcttggcagaattgtcccttcacatagtgtcatg	2419
Db	7834	aattgcctctacggggagcccccccgcttggcagaattgtcccttcacatagtgtgtcatg	7893
QY	2420	tgcgtcttccacagccttctctctcctcagctcacaagcgtcgcagtgtgtgtgagcagcag	2479
Db	7894	tgcgtcttccacagccttctctctcctcagctcacaagcgtcgcagtgtgtgtgagcagcag	7953
QY	2480	cgttggaaagcagctcccggaagcctactc--agggaggaacttgcgaattccaatgcct	2536
Db	7954	cgttggaaagcagctcccggaagcctactcaggaaggaagacttgcgaattcccaatgcct	8013
QY	2537	tccaagcttggagattcagctctactgtttagagcgcacacggttgcaggaac	2583
Db	8014	tccaagcttggagattcagctctactgtttagagcgcacacggttgcaggaac	8060
RESULT 12			
AAC67021	AAC67021 standard; DNA; 6076 bp.		
XX	AC	AAC67021;	
XX	DT	27-MAR-2001 (first entry)	
XX	DE	PERV env protein coding sequence SEQ ID NO: 21.	
XX	KW	Xenotransplantation; infectious agent; vaccine; ds.	
XX	OS	Porcine endogenous retrovirus.	
XX	PN	W0200071726-A1.	
XX	PD	30-NOV-2000.	
XX	PF	24-MAY-2000; 2000MO-US14296.	
XX	PR	24-MAY-1999; 99US-0135631.	
XX	PA	(MAYO-) MAYO MEDICAL VENTURES.	
XX	PI	Federspiel MJ;	
XX	RR	WPI; 2001-032041/04.	

[illegible]

```
Db 4348 atcgagcgtggtgcatatggtcgaagagcgtgcaaatgtacttggcagagata 4407
      |||
Qy 862 actgtaatatgcccccatgcatctggagaaacatcgtctggcagccccccagccagatgagt 921
      |||
Db 4408 actgttaatatgcccccatgcatctggagaaacatcgtctggcagccccccagccagatgagt 4467
      |||
Qy 922 accaaagcccgatgacccaatataaaagcgtctctcagagagaggtcactctgct 981
      |||
Db 4468 accaaagcccgatgacccaatataaaagcgtctctcagagagaggtcactctgct 4527
      |||
Qy 982 ccacacagccgctctcaaccctgcaactcttctgctgaagagactgtaaccagtgact 1041
      |||
Db 4528 ccacacagccgctctcaaccctgcaactcttctgctgaagagactgtaaccagtgact 4587
      |||
Qy 1042 catgattgcatcaactatgattgagagagcgtgggtccgcaagagaccttaacagacta 1101
      |||
Db 4588 catgattgcatcaactatgattgagagagcgtgggtccgcaagagaccttaacagacta 4647
      |||
Qy 1102 ccgctgactggagaaagtctaaacctgctcaactgagcaagacagctatgtgtgaaagt 1161
      |||
Db 4648 ccgctgactggagaaagtctaaacctgctcaactgagcaagacagctatgtgtgaaagt 4707
      |||
Qy 1162 aagagagatggtctggggtggcagctgtgtgaaaggagccgacagactctggcagacgctg 1221
      |||
Db 4708 aagagagatggtctggggtggcagctgtgtgaaaggagccgacagactctggcagacgctg 4767
      |||
Qy 1222 ccggaaggaacttaacgccaagagcgtgagctcaactgagccctcagcaagcgtctgctgctg 1281
      |||
Db 4768 ccggaaggaacttaacgccaagagcgtgagctcaactgagccctcagcaagcgtctgctgctg 4827
      |||
Qy 1282 gccgaagggaaatccataaactlatacagacagcaggtatgctcttgcagctgcacac 1341
      |||
Db 4828 gccgaagggaaatccataaactlatacagacagcaggtatgctcttgcagctgcacac 4887
      |||
Qy 1342 gtacacgggtgcatctataaacaaggggtgtgttacctcagcagagggaggaataaag 1401
      |||
Db 4888 gtacacgggtgcatctataaacaaggggtgtgttacctcagcagagggaggaataaag 4947
      |||
Qy 1402 aacaaaggggaatctctaagccttagabgacctatacttggccaanaagcgttagactatc 1461
      |||
Db 4948 aacaaaggggaatctctaagccttagabgacctatacttggccaanaagcgttagactatc 5007
      |||
Qy 1462 atacaactgtcctgacatcagaagaacgaagatcatalctataaggaacacagatgct 1521
      |||
Db 5008 atacaactgtcctgacatcagaagaacgaagatcatalctataaggaacacagatgct 5067
      |||
Qy 1522 gaccgggtgtgccaagcagcagccagcgtgttaacctctgctcatalatagaacgccc 1581
      |||
Db 5068 gaccgggtgtgccaagcagcagccagcgtgttaacctctgctcatalatagaacacccc 5127
      |||
Qy 1582 aaagccccaagaccagacagacagctagacagcgtgcaagagatgaagagataaagagata 1641
      |||
Db 5128 aaagccccaagaccagacagacagctagacagcgtgcaagagatgaagagataaagagata 5187
      |||
Qy 1642 gaccagttctctgagactccggaaggagcctgtatatacctcatatggaaggaatcctg 1701
      |||
Db 5188 gaccagttctctgagactccggaaggagcctgtatatacctcatatggaaggaatcctg 5247
      |||
Qy 1702 ccccaagaagaagggttagatatagttccaacagatatacgtcttaaccacctaagaact 1761
      |||
Db 5248 ccccaagaagaagggttagatatagttccaacagatatacgtcttaaccacctaagaact 5307
      |||
Qy 1762 aaacacactgacagcagttgtgcaagaacatccctcatatcatgttctggaagctcaagagagt 1821
      |||
Db 5308 aaacacactgacagcagttgtgcaagaacatccctcatatcatgttctggaagctcaagagagt 5367
      |||
Qy 1822 gctgacactcggtgtgcaaacatgtgtgacctgccaagctggttaagtatactctccaga 1881
      |||
Db 5368 gctgacactcggtgtgcaaacatgtgtgacctgccaagctggttaagtatactctccaga 5427
      |||
Qy 1882 atacctcaggaagaagactaaaggggagacccagcgctcaactgggaagtggacttc 1941
      |||
```

```
Db 5428 atacctcaggaagaagactaaaggggagacccagcgctcaactgggaagtggacttc 5487
      |||
Qy 1942 actgaagtaaaagccggttaataataggaacaaatatactatgtttttgtgaacacctt 2001
      |||
Db 5488 actgaagtaaaagccggttaataataggaacaaatatactatgtttttgtgaacacctt 5547
      |||
Qy 2002 tcaagatgggttagaggtcttccctcaatgaagaagacccctcaacccgtggtgctaaagaa 2061
      |||
Db 5548 tcaagatgggttagaggtcttccctcaatgaagaagacccctcaacccgtggtgctaaagaa 5607
      |||
Qy 2062 atactggaagaaattttccaaagatttgaaatacctaaggtlaaaggtlcaagaacatgct 2121
      |||
Db 5608 atactggaagaaattttccaaagatttgaaatacctaaggtlaaaggtlcaagaacatgct 5667
      |||
Qy 2122 ccagcttctgttcccaaggttaagtcaaggaactggtccaagaatattggagattgtgaa 2181
      |||
Db 5668 ccagcttctgttcccaaggttaagtcaaggaactggtccaagaatattggagattgtgaa 5727
      |||
Qy 2182 ctgcatgtgcatcaaaccccaaggtcaaggaactggtgaaggttaagaaataaccatt 2241
      |||
Db 5728 ctgcatgtgcatcaaaccccaaggtcaaggaactggtgaaggttaagaaataaccatt 5787
      |||
Qy 2242 aaagaagaccccttaacccaattgaccacagagactggtcatlaatgattgagctctcctg 2301
      |||
Db 5788 aaagaagaccccttaacccaattgaccacagagactggtcatlaatgattgagagctctcctg 5847
      |||
Qy 2302 ccccttgtgtttttttaggtgtgaagaaacccctggagacgtttgtggtcagccctatgaa 2361
      |||
Db 5848 ccccttgtgtttttttaggtgtgaagaaacccctggagacgtttgtggtcagccctatgaa 5907
      |||
Qy 2362 ttgcttaacggggagaccccccgttggcagaaattgaccttgcacatagttgtgtgtg 2421
      |||
Db 5908 ttgcttaacggggagaccccccgttggcagaaattgaccttgcacatagttgtgtgtg 5967
      |||
Qy 2422 ctgcttcccaagccttgttctcttaaggtcacaagcgtcgaagtgtgtgaagcagcagcg 2481
      |||
Db 5968 ctgcttcccaagccttgttctcttaaggtcacaagcgtcgaagtgtgtgaagcagcagcg 6027
      |||
Qy 2482 tggagacagctccggagagcctactcagagagag 2515
      |||
Db 6028 tggagacagctccggagagcctactcagagagag 6061
      |||
```

RESULT 13
AAC67032
ID AAC67032 standard; DNA; 6076 BP.
XX
AC AAC67032;
XX
DT 27-MAR-2001 (first entry)
XX
DE PERV env protein coding sequence SEQ ID NO: 32.
XX
KW Xenotransplantation; infectious agent; vaccine; ds.
XX
OS Porcine endogenous retrovirus.
PN WO200071726-A1.
XX
PD 30-NOV-2000.
XX
PF 24-MAY-2000; 2000WO-US14296.
XX
PR 24-MAY-1999; 99US-0135631.
XX
PA (MAYO-) MAYO MEDICAL VENTURES.
XX
PI Federspiel MJ;
XX
DR WPI; 2001-032041/04.
XX
PT Inhibiting or preventing infectious agent transmission in mammalian
transplant recipients, by introducing recombinant DNA comprising DNA

PT encoding extracellular proteins of the agent into donor cells, such as
PT swine cells -
XX
PS Claim 16, Page 117-119, 144pp, English.

XX
CC The present invention provides a method to prevent the transmission of
CC infectious agents during xenotransplantation. This involves introducing
CC to donor swine cells a recombinant DNA encoding a peptide fragment from
CC the infectious agent, and then introducing these cells into the
CC transplant recipient.
XX

Sequence 6076 BP; 1619 A; 1505 C; 1598 G; 1354 T; 0 other;

Query Match 72.7%; Score 2414; DB 22; Length 6076;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 2444; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 22 gatgctctctctgctgagatcacacccactagccacacattttgcttcgaatg 81
DB 3568 gatgctctctctgctgagatcacacccactagccacacattttgcttcgaatg 3627
QY 82 agagatccaggtacgaggaagaccgagctacacttgacccgactgcccagaagctc 141
DB 3628 agagatccaggtacgaggaagaccgagctacacttgacccgactgcccagaagctc 3687
QY 142 aagaactcccgacacatttgacgagccctacacagagaccctggccaaactcagatc 201
DB 3688 aagaactcccgacacatttgacgagccctacacagagaccctggccaaactcagatc 3747
QY 202 caacacccctcaggtgacccctcctcagttacgtgagatgacctgcttcgagggagccac 261
DB 3748 caacacccctcaggtgacccctcctcagttacgtgagatgacctgcttcgagggagccac 3807
QY 262 aaaaagagactgcttagaagtaagagagcactactgctggaattgtctgacagatc 321
DB 3808 aaaaagagactgcttagaagtaagagagcactactgctggaattgtctgacagatc 3867
QY 322 agagcctctgcttagaagagccagatttgagagagagtaacatacttggggtagatc 381
DB 3868 agagcctctgcttagaagagccagatttgagagagagtaacatacttggggtagatc 3927
QY 382 ttgctgggagcgagcagatgagtcgagagcgaggaagaaactgtatgccaagatac 441
DB 3928 ttgctgggagcgagcagatgagtcgagagcgaggaagaaactgtatgccaagatac 3987
QY 442 gcccacacacagcagcaaaagtgagagagtttttggggacagctgatttcagactg 501
DB 3988 gcccacacacagcagcaaaagtgagagagtttttggggacagctgatttcagactg 4047
QY 502 tggatcccgaggtttgagacatttagacgcccactctacccgcttaaccaaagaagag 561
DB 4048 tggatcccgaggtttgagacatttagacgcccactctacccgcttaaccaaagaagag 4107
QY 562 ggaattctctggtgctctgagcagcagagagcatttgatcaaaaagccctgctg 621
DB 4108 ggaattctctggtgctctgagcagcagagagcatttgatcaaaaagccctgctg 4167
QY 622 agcgcactgctgctgcccctccctgagcttaacaaaccttacccttcttgatgagag 681
DB 4168 agcgcactgctgctgcccctccctgagcttaacaaaccttacccttcttgatgagag 4227
QY 682 cgttagagagtagccgagagagttttaaccaaacccttagagacatgagagagacgtc 741
DB 4228 cgttagagagtagccgagagagttttaaccaaacccttagagacatgagagagacgtc 4287
QY 742 gctactctgtcaagaagagcttgatcctgtagccagtggtgtggccgtatgtcgtgaag 801
DB 4288 gctactctgtcaagaagagcttgatcctgtagccagtggtgtggccgtatgtcgtgaag 4347
QY 802 atcgcaactgtgacatactggttaagaagcgtgacaaatgtgacttgggacagatata 861
DB 4348 atcgcaactgtgacatactggttaagaagcgtgacaaatgtgacttgggacagatata 4407

QY 862 actgtaatagcccccatgcatctgagagaacatcgttcgcaagccccagacagatgag 921
DB 4408 actgtaatagcccccatgcatctgagagaacatcgttcgcaagccccagacagatgag 4467
QY 922 accaagcgcgcagatgacccacatcaaaagcgtctctcacaagaagagcttctgct 981
DB 4468 accaagcgcgcagatgacccacatcaaaagcgtctctcacaagaagagaggtacgtt 4527
QY 982 ccacagcgcgtctcacaagcgtctctcacaagaagagagaggtacgtt 1041
DB 4528 ccacagcgcgtctcacaagcgtctctcacaagaagagagaggtacgtt 4587
QY 1042 catgattgccaatcaactttagttagagagagctgggttcgcaagaagccttacagacata 1101
DB 4588 catgattgccaatcaactttagttagagagagctgggttcgcaagaagccttacagacata 4647
QY 1102 ccgctgacttgaggaagctgtaactggtctactcaacgaaacacatgtgtgtggaag 1161
DB 4648 ccgctgacttgaggaagctgtaactggtctactcaacgaaacacatgtgtgtggaag 4707
QY 1162 aagaagatggtctggggcgagctgtgtgagcggagcccgacgactggtggcagacgctg 1221
DB 4708 aagaagatggtctggggcgagctgtgtgagcggagcccgacgactggtggcagacgctg 4767
QY 1222 ccggaaggaacttcaagcgcaaaagcttgagctcaltggtccctcaagcaagcttgcgctg 1281
DB 4768 ccggaaggaacttcaagcgcaaaagcttgagctcaltggtccctcaagcaagcttgcgctg 4827
QY 1282 gcccgaaggaatccataacatctatcgagacgagagtgcttcttgagagcgccac 1341
DB 4828 gcccgaaggaatccataacatctatcgagacgagagtgcttcttgagagcgccac 4887
QY 1342 gtacacggtggccatctataaaacaaaggggtgtgttactcctaagcagggaggaataa 1401
DB 4888 gtacacggtggccatctataaaacaaaggggtgtgttactcctaagcagggaggaataa 4947
QY 1402 aaaaaaggaatctcctaagcctatctagaagccttaacttggccaaaaggtctgact 1461
DB 4948 aaaaaaggaatctcctaagcctatctagaagccttaacttggccaaaaggtctgact 5007
QY 1462 atacactgctctgagacatcagaagaaccaaagatccatctatctagaaggaacagatg 1521
DB 5008 atacactgctctgagacatcagaagaaccaaagatccatctatctagaaggaacagatg 5067
QY 1522 gacccggtctgcaaacagcagccagcagctgttlaaccttctgctcataatagaaagcc 1581
DB 5068 gacccggtctgcaaacagcagccagcagctgttlaaccttctgctcataatagaaagcc 5127
QY 1582 aaagcccaagaacccagagcagctacaccccttagaagagctggcagaagataaaaga 1641
DB 5128 aaagcccaagaacccagagcagctacaccccttagaagagctggcagaagataaaaga 5187
QY 1642 gacccggtctctgagacatcagaagaaccaaagatccatctatctagaaggaacagatg 1701
DB 5188 gacccggtctctgagacatcagaagaaccaaagatccatctatctagaaggaacagatg 5247
QY 1702 ccccaaaaagaaggttagaatactgccaacagatacctgcttaacccacttagaact 1761
DB 5248 ccccaaaaagaaggttagaatactgccaacagatacctgcttaacccacttagaact 5307
QY 1762 aaacactgcaagcagttggtcagaacatctccttcaatgtctgaagctacccaagatg 1821
DB 5308 aaacactgcaagcagttggtcagaacatctccttcaatgtctgaagctacccaagatg 5367
QY 1822 gctgactggtgtgtaaaacatgtgtgcccgtgagcagctgttaagcttaacccctcaga 1881
DB 5368 gctgactggtgtgtaaaacatgtgtgcccgtgagcagctgttaagcttaacccctcaga 5427
QY 1882 atacactcaggaagaagactaaagggaagccacccagcgctcactctggaagtgacttc 1941
DB 5428 atacactcaggaagaagactaaagggaagccacccagcgctcactctggaagtgacttc 5487

QY 1942 actgaggtaaagccggctaaatacagaaacataatctatgtgtttttagaacacctt 2001
 DB 5488 actgaggtaaagccggctaaatacagaaacataatctatgtgtttttagaacacctt 5547
 QY 2002 taagagtggtgaggtctatcctactaagaagaagactcaacccgtgtgttaagaa 2061
 DB 5548 taagagtggtgaggtctatcctactaagaagaagactcaacccgtgtgttaagaa 5607
 QY 2062 atactggaagaattttcccaagatttgaaactaaggaatgaatggtgtcagacaagt 2121
 DB 5608 atactggaagaattttcccaagatttgaaactaaggaatgaatggtgtcagacaagt 5667
 QY 2122 ccagcttcgtgtgccaggttaagtcaggaactgtgcccaagatactgggattgttgaa 2181
 DB 5668 ccagcttcgtgtgccaggttaagtcaggaactgtgcccaagatactgggattgttgaa 5727
 QY 2182 ctgcattgtgcatcacagaccacaaagctcagagcagtagagaggaatagaacat 2241
 DB 5728 ctgcattgtgcatcacagaccacaaagctcagagcagtagagaggaatagaacat 5787
 QY 2242 aagagacccttaaccaaatgacacagagactgtgcatatgtatgtgtgtctctctg 2301
 DB 5788 aagagacccttaaccaaatgacacagagactgtgcatatgtatgtgtgtctctctg 5847
 QY 2302 cccttgtgtcttttagaggttagagacacccctgtgacagttgtggtctacccctatga 2361
 DB 5848 cccttgtgtcttttagaggttagagacacccctgtgacagttgtggtctacccctatga 5907
 QY 2362 ttgctctacgggggagaccccccgtgtgacagaaattgtccttgcacatagtgtcgtatg 2421
 DB 5908 ttactctacgggggagaccccccgtgtgacagaaattgtccttgcacatagtgtcgtatg 5967
 QY 2422 ctgcttcccgagccttcttcttagaggtcagagcgtcgtgtgtgtgagcagcagcg 2481
 DB 5968 ctgcttcccgagccttcttcttagaggtcagagcgtcgtgtgtgtgagcagcagcg 6027
 QY 2482 tggaaagcagctccggagagcctactcagagagag 2515
 DB 6028 tggaaagcagctccggagagcctactcagagagag 6061

RESULT 14
 AAF77726
 ID AAF77726 standard; DNA; 7333 BP.
 AC AAF77726;
 XX 23-MAY-2001 (first entry)
 DT Defective retroviral genome isolated from PK-15 cell line.
 DE
 XX
 KW Retrovirus; graft transplantation; xenotransplantation; PK-15 cell line;
 ss.
 XX Unidentified.
 OS US6190861-B1.
 PN 20-FEB-2001.
 PD 13-DEC-1996; 96US-0766528.
 PE 14-DEC-1995; 95US-0572645.
 PR (GEHO) GEN HOSPITAL CORP.
 XX
 PA Fishman JA;
 FI
 XX WPI: 2001-256211/26.
 DR P-PSDB: AAB73282, AAB73283, AAB73284.
 XX Assessing risk of endogenous retroviruses in clinical practice and in
 PT xenotransplantation, comprises using probe sequences derived from swine

PM or miniature swine retroviral genome -
 XX
 PS Claim 1; Fig 2; 127pp; English.
 CC The present invention relates to a method for screening a cell or tissue
 CC for the presence or expression of a retrovirus (RV), comprising
 CC contacting a target nucleic acid from the cell or tissue with a second
 CC nucleic acid from the present invention (e.g. the present sequence or a
 CC fragment thereof). The method is useful for RV detection and to assess
 CC graft transplantation risk. Screening of animals allows the elimination
 CC of donors with active replication of known viruses. Inactive proviruses
 CC can be detected and inactivated, allowing identification and elimination
 CC of potential human pathogens derived from swine in a manner not possible
 CC in the outbred human organ donor population and is important to the
 CC development of human xenotransplantation.
 CC
 XX
 SQ Sequence 7333 BP; 1984 A; 1788 C; 1885 G; 1676 T; 0 other;

Query Match 53.4%; Score 1771.8; DB 22; Length 7333;
 Best Local Similarity 95.1%; Pred. No. 0;
 Matches 1871; Conservative 0; Mismatches 17; Indels 79; Gaps 1;

QY 1351 gccatctataacaaagggtgtgtgtctactcagcagggagggaaataaagaacaaagag 1410
 DB 3504 gctacccaagaagaaagggtgtgtgtctactcagcagggagggaaataaagaacaaagag 3563
 QY 1411 gaattctaaagcctatagaagccttactatcttccaaaaggctatatacactgt 1470
 DB 3564 gaattctaaagcctatagaagccttactatcttccaaaaggctatatacactgt 3623
 QY 1471 ccttgaatcagaagaagcgaagatctcatatctagaagggaacagatgtgtcagcggtt 1530
 DB 3624 ccttgaatcagaagaagcgaagatctcatatctagaagggaacagatgtgtcagcggtt 3683
 QY 1531 gcccaagaggcagccagcaggtgtgttaacctcttccataatagaacgcaccaagccca 1590
 DB 3684 gcccaagaggcagccagcaggtgtgttaacctcttccataatagaacgcaccaagccca 3743
 QY 1591 gaaccagacgacagatcacacctagaaagactgtgcaagagataaagaatagaccagttc 1650
 DB 3744 gaaccagacgacagatcacacctagaaagactgtgcaagagataaagaatagaccagttc 3803
 QY 1651 tctgagactccggagggagactgtctatacctatattggaagaaatcctgtcccacaa 1710
 DB 3804 tctgagactccggagggagactgtctatacctatattggaagaaatcctgtcccacaa 3863
 QY 1711 gaagggttaaatatgttccacagatcacatgctcaaccacctaaggaactaaacagctg 1770
 DB 3864 gaagggttaaatatgttccacagatcacatgctcaaccacctaaggaactaaacagctg 3923
 QY 1771 cagcagttgttcagaaatcccttatactatgttctgaagctacacagaggtgtgtactcg 1830
 DB 3924 cagcagttgttcagaaatcccttatactatgttctgaagctacacagaggtgtgtactcg 3983
 QY 1831 gtgttaaacattgtgtgcccgtgcagctgtgttaatgctcaatccctccaagaatacctca 1890
 DB 3984 gtgttaaacattgtgtgcccgtgcagctgtgttaatgctcaatccctccaagaatacctca 4043
 QY 1891 ggaagaagactaaagggaagccacccagcgcctacattgtggaagtgtgtactcactagta 1950
 DB 4044 ggaagaagactaaagggaagccacccagcgcctacattgtggaagtgtgtactcactagta 4103
 QY 1951 aagcggctaaatacggaaacaatactatgtgttttctagacacctttcagagatgg 2010
 DB 4104 aagcggctaaatacggaaacaatactatgtgttttctagacacctttcagagatgg 4163
 QY 2011 gttagagcttactcctactaagaagaagacttcaacccgtgtgtgtcgaagaatactggag 2070
 DB 4164 gttagagcttactcctactaagaagaagacttcaacccgtgtgtgtcgaagaatactggag 4223
 QY 2071 gaaatttccaaagattggaatacctaaggtcaatagaggtcagaacatgtgtcagcttc 2130

Oy	3217	gcaacagcgggtccaaaaagtctaccgaataagaaccaatgaacttcattcgtagacct	3270'
Oy	3218		
Db	5285	gcacaagcgggtccaaaaagtacgcgaataaagcacaaataagctgcattcgtagacct	5344'
Oy	3271	agatacttaaaaaataagtttcctaataaaaaaaaaaaaaaaaaaaa	3317
Oy	3272		
Db	5345	agatacttaaaaaataagtttcacttgtaaaaggaaacagaanaata	5391
<hr/>			
RESULT#	15		
AAT74883	ID	AAT74883 standard; cDNA; 7393 BP.	
xx	AA74883:		
xx	AC		
xx	DT	09-FEB-1998 (first entry)	
xx	DE	Porcine retrovirus cDNA (defective).	
xx	KM	Retrovirus; Porcine; GAG protein; POL protein; ENV protein;	
xx	KW	Xenotransplantation; infectious; provirus; organ transplant; donor;	
xx	OS	activated virus; PCR; ss.	
xx	OS	Porcine retrovirus.	
FH	Key	Location/Qualifiers	
FT	CDS	598..2172	
FT		/**tag= a	
FT	mat_peptide	598..2169	
FT		/**tag= b	
FT		/note= "putative GAG protein"	
FT	CDS	2320..4737	
FT		/**tag= c	
FT		/note= "putative POL coding region (partial) as described in the specification"	
FT	mat_peptide	2320..3522	
FT		/**tag= d	
FT		/note= "putative POL protein (partial)"	
FT	mat_peptide	3516..4328	
FT		/**tag= e	
FT		/note= "putative POL protein (partial)"	
FT	mat_peptide	4332..4748	
FT		/**tag= f	
FT	CDS	4738..6725	
FT		/**tag= g	
FT		/note= "putative ENV coding region (partial) as described in the specification"	
FT	mat_peptide	4752..6722	
FT		/**tag= h	
FT		/note= "ENV protein (partial)"	
PN	WO9721836-A1.		
PD	19-JUN-1997.		
PF	13-DEC-1996;	96WO-US19660.	
PR	14-DEC-1995;	95US-0572645.	
PA	(GENO) GEN HOSPITAL CORP.		
PI	Fishman JA;		
DR	WIPI: 1997-332804/30.		
DR	P-PSDB: AAM32091-W32095.		
PT	New nucleic acid from porcine retro.viruses - used for detecting viruses in transplant or other tissue and for assessing risk of transmitting infection to graft recipient		
PS	Claim 16; Fig 2; 128bp; English.		


```

Oy 3151 tgltaacaatcctaccagttataatcaatttaattatggccatgggagatggaagaattg 3210
    |||||||
Db 5225 tgltaacaatcctaccagttataatcaatttaattatggccatgggagatggaagaattg 5284
    |||||||
Oy 3211 gcaacagcgggtacaaaaagatgtacgaaataaagcaataaagctgtcattcgttagaact 3270
    |||||||
Db 5285 gcaacagcgggtacaaaaagatgtacgaaataaagcaataaagctgtcattcgttagaact 5344
    |||||||
Oy 3271 agattacttaaaaataagttcacttaaaaaaataaaaaaata 3317
    |||||||
Db 5345 agattacttaaaaataagttcactgaaaaaagaaacagaata 5391
    |||||||

```

Search completed: February 24, 2002, 02:48:16
 Job time: 18462 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2002, 02:48:16 ; Search time 624.99 Seconds
(without alignments) 11260.633 Million cell updates/sec

Title: US-09-171-553B-3

Perfect score: 8209
Sequence: 1 gtgtgttcgcgcctgtgtgccc.....aaaaaaaaaaaaaaaaaaaaa 8209

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
1: N_Geneseq_1101: *
2: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT: *
3: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT: *
4: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT: *
5: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT: *
6: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT: *
7: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT: *
8: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT: *
9: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT: *
10: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT: *
11: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT: *
12: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT: *
13: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT: *
14: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT: *
15: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT: *
16: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT: *
17: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT: *
18: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT: *
19: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT: *
20: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT: *
21: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT: *
22: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT: *
23: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8209	100.0	8209	18	AAV09700 Porcine retrovirus
2	8171.8	99.5	8196	18	AAV09699 Porcine retrovirus
3	6282.4	76.5	8132	22	AAV7727 Nucleotide sequenc
4	6141	74.8	7892	18	AAV74884 Miniature swine re
5	5637.2	68.7	7873	22	AAV77023 PERV env protein c
6	5528.2	67.3	7362	22	AAV67019 PERV env protein c
7	5489.8	66.9	7333	22	AAV7726 Defective retrovir
8	5416.6	66.0	7393	18	AAV74883 Porcine retrovirus
9	5025.8	61.2	6076	22	AAV67021 PERV env protein c
10	5025.8	61.2	6076	22	AAV67032 PERV env protein c
11	4986	60.7	8060	18	AAV74811 Porcine retrovirus

12	4986	60.7	8060	22	AAV77725
13	3755.4	45.7	4402	22	AAV67020
14	3376.6	41.1	3482	22	AAV62749
15	3295.2	40.1	4918	22	AAV67022
16	32816	35.5	3320	18	AAV09698
17	2295.8	28.0	8655	20	AAV69750
18	2244.4	27.3	8088	21	AAZ45540
19	2234.4	27.2	8535	15	AAV73731
20	1941.6	23.7	1974	18	AAV09703
21	1747	21.3	8323	16	AAV081042
22	1747	21.3	8323	18	AAV80058
23	1747	21.3	8323	18	AAV05061
24	1737.4	21.2	8202	17	AAV09280
25	1735.6	21.1	8323	15	AAV058700
26	1700	20.7	8278	22	AAV08138
27	1697.8	20.7	8440	17	AAV13265
28	1696.6	20.7	8892	17	AAV04266
29	1687.4	20.6	8332	18	AAV05061
30	1687.4	20.6	8332	19	AAV42715
31	1687.4	20.6	8332	19	AAV19551
32	1687.4	20.6	8332	19	AAV15346
33	1687.4	20.6	8332	20	AAV28143
34	1687.4	20.6	8332	21	AAV40297
35	1687.4	20.6	8332	21	AAZ90064
36	1687.4	20.6	8332	21	AAZ45539
37	1686	20.5	8332	16	AAV08452
38	1684.2	20.5	8332	17	AAV07374
39	1680.4	20.5	8330	20	AAV77178
40	1596	19.4	10367	18	AAV80057
41	1596	19.4	10367	20	AAV07067
42	1576.8	19.2	10367	15	AAV058701
43	1568.4	19.1	2462	20	AAV82748
44	1550.4	18.9	7308	18	AAV90693
45	1550.4	18.9	7616	18	AAV90688

ALIGNMENTS

RESULT 1	AAV09700	standard; DNA; 8209 BP.
ID	AAV09700	standard; DNA; 8209 BP.
XX	AAV09700;	
DE	20-MAY-1998	(first entry)
XX	Porcine retrovirus DNA encoding the LTR and GAG, POL and ENV proteins.	
XX	Porcine retrovirus: POEV; POL protein; ENV protein; GAG protein;	
KW	virion core polypeptide; polymerase protein; envelope protein;	
KW	vacine; diagnosis; xenotransplantation; prophylactic; therapeutic; ds.	
XX	Porcine retrovirus.	
OS	Porcine retrovirus.	
XX	Key	Location/Qualifiers
FT	FT msc-feature	1..61
FT	FT msc-feature	/tag- a
FT	FT msc-feature	/note- "R LTR domain"
FT	FT msc-feature	1..588
FT	FT msc-feature	/tag- b
FT	FT msc-feature	62..143
FT	FT msc-feature	/tag- c
FT	FT msc-feature	/note- "U5 LTR domain"
FT	FT msc-feature	588..2162
FT	FT msc-feature	/tag- d
FT	FT msc-feature	/product= GAG protein
FT	FT msc-feature	/note- "viral core polypeptide"
FT	FT msc-feature	2163..5747
FT	FT msc-feature	/tag- e
FT	FT msc-feature	/product= POL protein
FT	FT msc-feature	/note- "polymerase peptide"
FT	FT msc-feature	5620..7590
FT	FT msc-feature	CDS

Tsukuba-1 cDNA.
PERV env protein c
Pig endogenous ret
PERV env protein c
Porcine retrovirus
Mus dunni endogeno
Complete nucleotid
GALV SEA70 genome.
Porcine retrovirus
Friend murine leuk
LTR clone of FB29
Friend ecotropic m
Novel AMP/MCF Viru
LTR clone of FB29
Canine retrovirus
Fely F6A provirus
Osteoinductive ret
Moloney murine leu
Moloney murine leu
Retroviral vector
Retroviral gag gen
MMLV gag-pol gene
Moloney murine leu
Complete nucleotid
MMLV gag/pol gene
MMLV gag/pol gene
MMLV gene sequenc
Sequence of plasmid
Glycopeptide expe
pURB332. Syntheti
Pig endogenous ret
Plasmid hCMV+Intro
Plasmid Ceb coding

FT	/tag= f
FT	/product= ENV protein
FT	/note= "envelope protein"
FT	7591..8209
FT	/tag= g
FT	7638..8106
FT	/tag= h
FT	/note= "u3 LTR domain"
FT	7643..7686
FT	/tag= i
FT	/note= "Binding site for transcription factors
FT	OCT-1, c-Myb, LfF-1, E47 and ETS-1"
FT	7719..7739
FT	/tag= j
FT	/note= "Binding site for transcription factors
FT	AP-4 and NF1"
FT	7744..7754
FT	/tag= k
FT	/note= "Binding site for transcription factor AP-1/TR"
FT	7773..7785
FT	/tag= l
FT	7793..7834
FT	/note= "Binding site for transcription factor ETS-1/GATA"
FT	/tag= m
FT	/note= "Binding site for transcription factors
FT	GATA, ETS-1, c-Myb, AP-1, GATA, E47"
FT	7841..7851
FT	/tag= n
FT	/note= "Binding site for transcription factor AP-1"
FT	7873..7883
FT	/tag= o
FT	/note= "Binding site for transcription factor AP-1"
FT	7885..7914
FT	/tag= p
FT	/note= "Binding site for transcription factors
FT	AP-1/GATA, c-Myb and AP-1"
FT	7928..7953
FT	/tag= q
FT	/note= "Binding site for transcription factors E47
FT	and ETS-1"
FT	7964..7968
FT	/tag= r
FT	7985..7997
FT	/tag= s
FT	/note= "Binding site for transcription factor GATA"
FT	8003..8007
FT	/tag= t
FT	8024..8036
FT	/tag= u
FT	/note= "Binding site for transcription factor GATA"
FT	8042..8046
FT	/tag= v
FT	8049..8060
FT	/tag= w
FT	/note= "Binding site for transcription factor
FT	AP-1/CREB"
FT	8075..8090
FT	/tag= x
FT	1..61
FT	/tag= y
FT	/note= "R LTR domain"
FT	8127..8135
FT	/tag= z
FT	/note= "PolyA downstream element"
FT	8161..8166
FT	/tag= aa
XX	
XX	W09740167-A1.
XX	
XX	30-OCT-1997.
XX	
XX	18-APR-1997;
XX	97WO-GB01087.

PR	10-FEB-1997:	97GB-0002668.
PR	19-APR-1996:	96GB-0008164.
XX		
PA	(IMUT-) IMUTRAN LTD.	
PA	(QONE-) Q-ONE BIOTECH LTD.	
XX		
PI	Galbraith DN, Haworth C, Lees GM, Smith KT,	
XX		
DR	WP1, 1997-535851/49.	
DR	P-PSDB: AAW39271; AAW39272; AAW39273.	
XX		
PT	Polynucleotide encoding porcine retrovirus expression product -	
PT	useful to develop products for use in vaccines, diagnosis and	
XX	xeno-transplantation	
XX		
PS	Claim 4; Fig 3; 69pp; English.	
XX		
CC	This DNA sequence encodes the porcine retrovirus (POEV) virion core	
CC	polypeptide (GAG), polymerase (POL) and envelope (ENV) proteins and	
CC	also includes the Long Terminal Repeat (LTR). These proteins can be used	
CC	to develop viral vaccines, antisense nucleic acids, ribozymes and other	
CC	antiviral agents. They can also be used in xeno-transplantation	
CC	technology and as diagnostic tools.	
XX		
XX		
SO	Sequence 8209 BP; 2168 A; 2064 C; 2154 G; 1823 T; 0 other;	

Query Match	100.0%;	Score 8209;	DB 18;	Length 8209;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 8209; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0

QY	1	gtgagtaagactgtgaggcccaagacgcttggaataaaatctctgtcgtttcgtac	60
Db	1	gtgagtaagactgtgaggcccaagacgcttggaataaaatctctgtcgtttcgtac	60
QY	61	aagacgcgtctctgtatgatatttgagggtgtgcgcctctccgagcccgagagggat	120
Db	61	aagacgcgtctctctgtatgatatttgagggtgtgcgcctctccgagcccgagagggat	120
QY	121	tgtcttttaactgagccttcactttgggtgcgttgagccggaaatctctgcagaccctta	180
Db	121	tgtcttttaactgagccttcactttgggtgcgttgagccggaaatctctgcagaccctta	180
QY	181	caaccggagaaaccgacttggagagtaaaaggatccccctttggaaacgtgtgtgtgcgcgc	240
Db	181	caaccggagaaaccgacttggagagtaaaaggatccccctttggaaacgtgtgtgtgcgcgc	240
QY	241	ggcgcctctgtctcgaagtgatctgtcttccggagatgcgcgccttcgggtttcagaatgctct	300
Db	241	ggcgcctctgtctcgaagtgatctgtcttccggagatgcgcgccttcgggtttcagaatgctct	300
QY	301	ctcagaaccgtaaagactgtgagagactgtatcaagacagacgtgtctagagagatacaagagctg	360
Db	301	ctcagaaccgtaaagactgtgagagactgtatcaagacagacgtgtctagagagatacaagagctg	360
QY	361	ccaaccctggagagacgcccccgagagtgaggagaaacccagagacgcgtgggtgtctctact	420
Db	361	ccaaccctggagagacgcccccgagagtgaggagaaacccagagacgcgtgggtgtctctact	420
QY	421	gtcgcgtcaagagacgagatctgtgtgttgtaagcgaagcttccccctccgcgcgtccg	480
Db	421	gtcgcgtcaagagacgagatctgtgtgttgtaagcgaagcttccccctccgcgcgtccg	480
QY	481	actctttgacctgtgttggaagaacggagcgggtcgcggtgtctcgatctgtgttt	540
Db	481	actctttgacctgtgttggaagaacggagcgggtcgcggtgtctcgatctgtgttt	540
QY	541	ctgtctcgtgtctctgtctctgtgtgcgccttctcaagtttaatacgtggaacagag	600
Db	541	ctgtctcgtgtgtctctgtctctgtgtgcgccttctcaagtttaatacgtggaacagag	600
QY	601	tgatacccccttagtttgactctcgacatctggactgaagttagatcccaaggctcata	660
Db	601	tgatacccccttagtttgactctcgacatctggactgaagttagatcccaaggctcata	660

Db 601 tgactaccccccttagtttgactctcgaccattgagctggaagttagatccaagggctcata 660
QY 661 attgttcagttcaaggttaagaaggaccttggcagagcttctgtgctcttgaatgcca 720
Db 661 attgttcagttcaaggttaagaaggaccttggcagagcttctgtgctcttgaatgcca 720
QY 721 cattcgatgttgatgagccatcagaagggaaccttaattcgaataatctctgtgtta 780
Db 721 cattcgatgttgatgagccatcagaagggaaccttaattcgaataatctctgtgtta 780
QY 781 aggcacatcatttttcagactgagccgctctcattccgatacagaagccctatctctta 840
Db 781 aggcacatcatttttcagactgagccgctctcattccgatacagaagccctatctctta 840
QY 841 cgttgcaagatcttggcagaagaatctccgcgcatgtgttaaacctcttgctaataacaa 900
Db 841 cgttgcaagatcttggcagaagaatctccgcgcatgtgttaaacctcttgctaataacaa 900
QY 901 gaaagccaggtcccggaatcctgtgctcttgagagaaanaaacactcggccgnaaag 960
Db 901 gaaagccaggtcccggaatcctgtgctcttgagagaaanaaacactcggccgnaaag 960
QY 961 tcgagccctctctctgfatctaccccgagatcgagagcgccgagcttggccgnaacc 1020
Db 961 tcgagccctctctctgfatctaccccgagatcgagagcgccgagcttggccgnaacc 1020
QY 1021 caactgttcccccaaccccttaccagacagagtgctgtgaaggagaccttccccctc 1080
Db 1021 caactgttcccccaaccccttaccagacagagtgctgtgaaggagaccttccccctc 1080
QY 1081 ctgagagctcgggtgtgagaggagacctgtcgcggagactcggaagccgagagcgccacc 1140
Db 1081 ctgagagctcgggtgtgagaggagacctgtcgcggagactcggaagccgagagcgccacc 1140
QY 1141 cggagcgagacagagagatcgagatattaccgctgcgcactatagccctcccatgtgca 1200
Db 1141 cggagcgagacagagagatcgagatattaccgctgcgcactatagccctcccatgtgca 1200
QY 1201 ggggccaattgacagccctccagatctgagcccttctctgcagactcctatattgtga 1260
Db 1201 ggggccaattgacagccctccagatctgagcccttctctgcagactcctatattgtga 1260
QY 1261 aaactaacatccccccttctcgaagatcccccaacgctcgaaggtgtgtgaggtccc 1320
Db 1261 aaactaacatccccccttctcgaagatcccccaacgctcgaaggtgtgtgaggtccc 1320
QY 1321 ttatgttctctcaccagctacttggatgtatgttcaacgctgtcgcacacacttcta 1380
Db 1321 ttatgttctctcaccagctacttggatgtatgttcaacagctgtcgcacacacttcta 1380
QY 1381 caaccgagagcagagagaaattctgtctagaagctagaanaaaatgttccctggggcgag 1440
Db 1381 caaccgagagcagagagaaattctgtctagaagctagaanaaaatgttccctggggcgag 1440
QY 1441 ggcgacccacagcttgcanaatgagatlgacatlggattcccttgactcgcccggt 1500
Db 1441 ggcgacccacagcttgcanaatgagatlgacatlggattcccttgactcgcccggt 1500
QY 1501 gggactacaacacagctggaaggttagagagcttgaanaactatccgccaagctctgttg 1560
Db 1501 gggactacaacacagctggaaggttagagagcttgaanaactatccgccaagctctgttg 1560
QY 1561 cgggtctccggcgccctcaagacgcccactaatttgcttaaggttagagaggtgagtc 1620
Db 1561 cgggtctccggcgccctcaagacgcccactaatttgcttaaggttagagaggtgagtc 1620
QY 1621 agggacggaacgaacctccctcgttattcttgaagagctcatggaagccttcaagcggt 1680
Db 1621 agggacggaacgaacctccctcgttattcttgaagagctcatggaagccttcaagcggt 1680
QY 1681 tcaccccttttgatctacacctcagaagccagaagactaagtgccctgtctattgt 1740
Db 1681 tcaccccttttgatctacacctcagaagccagaagactaagtgccctgtctattgt 1740

QY 1741 ggcagtcgggctctgatatcagaagaaccttcagagactggaaaggtttacagaggtcg 1800
Db 1741 ggcagtcgggctctgatatcagaagaaccttcagagactggaaaggtttacagaggtcg 1800
QY 1801 agtlaacgttactagtgagagagcgagagaaggtgtatctacacgaagggagacagaag 1860
Db 1801 agtlaacgttactagtgagagagcgagagaaggtgtatctacacgaagggagagag 1860
QY 1861 agagaagacagagaagaagaagaagagaagaagaagaggaagacgttgatagcgyc 1920
Db 1861 agagaagacagagaagaagaagaagagaagaagaagaggaagacgttgatagcgyc 1920
QY 1921 aagagaagaatttgaactaagatcttggccgagtggttgaagggaaagacagaggaga 1980
Db 1921 aagagaagaatttgaactaagatcttggccgagtggttgaagggaaagacagaggaga 1980
QY 1981 gagaagagagattttagaanaattaggtcagggccctagacagtcagaggacctgggcaata 2040
Db 1981 gagaagagagattttagaanaattaggtcagggccctagacagtcagaggacctgggcaata 2040
QY 2041 ggaacccactcgaacagagacagtggtcglatgtgaagaaagaacactgggcaagga 2100
Db 2041 ggaacccactcgaacagagacagtggtcglatgtgaagaaagaacactgggcaagga 2100
QY 2101 actgccccaaagagaggaacaaagacagcgagcttctatgtagaagaataaagatt 2160
Db 2101 actgccccaaagagaggaacaaagacagcgagcttctatgtagaagaataaagatt 2160
QY 2161 agggagagaggggtctcgagacccctcccgagcccaaggstaacttgaaggttgaagggc 2220
Db 2161 agggagagaggggtctcgagacccctcccgagcccaaggstaacttgaaggttgaagggc 2220
QY 2221 aaccagttagttcctggtttgataccgagagagagatcagtgctgcataacattag 2280
Db 2221 aaccagttagttcctggtttgataccgagagagagatcagtgctgcataacattag 2280
QY 2281 gaaacataaagaagaanaaaatccctgggtatgtgtcccaagggcacaagcgatccat 2340
Db 2281 gaaacataaagaagaanaaaatccctgggtatgtgtcccaagggcacaagcgatccat 2340
QY 2341 ggaactaccggaagacgcttgacttggagtgaggaggglaaaccaactggtttctgtca 2400
Db 2341 ggaactaccggaagacgcttgacttggagtgaggaggglaaaccaactggtttctgtca 2400
QY 2401 ttccttagtgcccaagraccccttctagttagagacttactgacccaagatggaggtccaa 2460
Db 2401 ttccttagtgcccaagraccccttctagttagagacttactgacccaagatggaggtccaa 2460
QY 2461 ttctcttgaacagaagaagacagaagtgctgtgataaacaacccaactggttga 2520
Db 2461 ttctcttgaacagaagaagacagaagtgctgtgataaacaacccaactggttga 2520
QY 2521 cccctcaattagatgataatgcactatctccccaagtaaagctgatacaagata 2580
Db 2521 cccctcaattagatgataatgcactatctccccaagtaaagctgatacaagata 2580
QY 2581 taagatcccggttggagcggttccccaagcttggcagaacccggaaggtgtgttg 2640
Db 2581 taagatcccggttggagcggttccccaagcttggcagaacccggaaggtgtgttg 2640
QY 2641 caaagcaagtcccccacaggttatctcaactgaagccaggtgctacacagtatcagta 2700
Db 2641 caaagcaagtcccccacaggttatctcaactgaagccaggtgctacacagtatcagta 2700
QY 2701 gacagtacccttgattagagaggtcgaagaagaaatttggccgcagatgttcaagaattaa 2760
Db 2701 gacagtacccttgattagagaggtcgaagaagaaatttggccgcagatgttcaagaattaa 2760
QY 2761 tccaacagggcatcaggttccgttccaatcccttggaatactccctgtgaccggtta 2820
Db 2761 tccaacagggcatcaggttccgttccaatcccttggaatactccctgtgaccggtta 2820

QY 2821 ggaagcctgggacaaatgatataatgacagctacagagcttgagagaggtcaataaagg 2880
DB 2821 ggaagcctgggacaaatgatataatgacagctacagagcttgagagaggtcaataaagg 2880
QY 2881 tgcagagacatacaaccaacaggtcccgaaaccttaaacctcttagagccctccgcctg 2940
DB 2881 tgcagagacatacaaccaacaggtcccgaaaccttaaacctcttagagccctccgcctg 2940
QY 2941 aacggaactggtlaacagatltggaacttaaaagatgcctctctctgctgagatlaacc 3000
DB 2941 aacggaactggtlaacagatltggaacttaaaagatgcctctctctgctgagatlaacc 3000
QY 3001 ccaactagccaaacctttctgcttgatgagatgagatccaggtlaacggaagacccggc 3060
DB 3001 ccaactagccaaacctttctgcttgatgagatgagatccaggtlaacggaagacccggc 3060
QY 3061 agctcaactggagccagactgcccgaaggttcaagaactcccgcgaacctcttagcaag 3120
DB 3061 agctcaactggagccagactgcccgaaggttcaagaactcccgcgaacctcttagcaag 3120
QY 3121 ccctacacagagagacctgccaacttcagatccaaacctcaagtgacctctccagt 3180
DB 3121 ccctacacagagagacctgccaacttcagatccaaacctcaagtgacctctccagt 3180
QY 3181 acgtgagatgacctgtctctgctggagagcccaaaacagagactgcttagaagatlaag 3240
DB 3181 acgtgagatgacctgtctctgctggagagcccaaaacagagactgcttagaagatlaag 3240
QY 3241 cactactgctggaattgtctgacctagctacagagcctctgctcaaaagagcccaattc 3300
DB 3241 cactactgctggaattgtctgacctagctacagagcctctgctcaaaagagcccaattc 3300
QY 3301 gcaagagagaggttaacatacttgggtacagattgagggcgagcgatgctgacag 3360
DB 3301 gcaagagagaggttaacatacttgggtacagattgagggcgagcgatgctgacag 3360
QY 3361 aggcacggaagaagaactgtatccagaataccggcccaacacagcccaaaacaggtgag 3420
DB 3361 aggcacggaagaagaactgtatccagaataccggcccaacacagcccaaaacaggtgag 3420
QY 3421 agtttttgggacagctgagatttgcagacttgcagactccggggttgcacacttaag 3480
DB 3421 agtttttgggacagctgagatttgcagacttgcagactccggggttgcacacttaag 3480
QY 3481 ccccaactacccgctacccaagaagaaggggatctctctggtcctctgagacacaga 3540
DB 3481 ccccaactacccgctacccaagaagaaggggatctctctggtcctctgagacacaga 3540
QY 3541 aggcatttgatgctataaaaaagccctgtgtgagcgacactgtcttgcctcccgagc 3600
DB 3541 aggcatttgatgctataaaaaagccctgtgtgagcgacactgtcttgcctcccgagc 3600
QY 3601 taactaaaccttacccttatgtgataagcgttaagggagtagcccgagaggtttta 3660
DB 3601 taactaaaccttacccttatgtgataagcgttaagggagtagcccgagaggtttta 3660
QY 3661 cccaacccctagagcaatgagagagacctgttgcctacactgtcaaaagagcttgacctg 3720
DB 3661 cccaacccctagagcaatgagagagacctgttgcctacactgtcaaaagagcttgacctg 3720
QY 3721 taagcagttgttgcccgctatgtctgaagcgtatgcagctgtgagcactatgttcaag 3780
DB 3721 taagcagttgttgcccgctatgtctgaagcgtatgcagctgtgagcactatgttcaag 3780
QY 3781 acgctgcaaatgtgacttggagacagataatactgtaatagcccccattgacattgaga 3840
DB 3781 acgctgcaaatgtgacttggagacagataatactgtaatagcccccattgacattgaga 3840
QY 3841 acatgcttcggagccccagacagatgagatgacaaagcccgatgacacacatataca 3900
DB 3841 acatgcttcggagccccagacagatgagatgacaaagcccgatgacacacatataca 3900
QY 3901 gctgcttctccacagagaggtcacttctgctccacacgctcctcaacctgccaactc 3960

DB 3901 gctgcttctccacagagaggtcacttctgctccacacgctcctcaacctgccaactc 3960
QY 3961 ttctgcttgaagagactgtaaacagtgactatcatgattgacataactatgattgag 4020
DB 3961 ttctgcttgaagagactgtaaacagtgactatcatgattgacataactatgattgag 4020
QY 4021 agactgggttcgcaagaccttacagacataccgcttgactgagagagtgtaactggt 4080
DB 4021 agactgggttcgcaagaccttacagacataccgcttgactgagagagtgtaactggt 4080
QY 4081 tcactacggaagcaactatgtgtgagaggttaagagagtgctggcgagtggtg 4140
DB 4081 tcactacggaagcaactatgtgtgagaggttaagagagtgctggcgagtggtg 4140
QY 4141 acgggacccgacagatctggccagagcctccggaagaagacttaagcgcaaaagcg 4200
DB 4141 acgggacccgacagatctggccagagcctccggaagaagacttaagcgcaaaagcg 4200
QY 4201 agctatggccctacgcaagcttggcgctggccgaagaggaatccataaacttata 4260
DB 4201 agctatggccctacgcaagcttggcgctggccgaagaggaatccataaacttata 4260
QY 4261 cggagacaggtatgtccttggcactgacacagtaacggggccatctataaagaag 4320
DB 4261 cggagacaggtatgtccttggcactgacacagtaacggggccatctataaagaag 4320
QY 4321 ggttgccttaacctgaagaggaagaataaagaagaagaagaaattcaagccattag 4380
DB 4321 ggttgccttaacctgaagaggaagaataaagaagaagaagaaattcaagccattag 4380
QY 4381 aagccttaacttgcgaaaaaagctagctatatacaactgtctcgtgacatcaagaagca 4440
DB 4381 aagccttaacttgcgaaaaaagctagctatatacaactgtctcgtgacatcaagaagca 4440
QY 4441 aagatctcatctagaaggaacagatgctgcagcgggttgcgaagcagagccagg 4500
DB 4441 aagatctcatctagaaggaacagatgctgcagcgggttgcgaagcagagccagg 4500
QY 4501 ctgttaaccttgccttaataaagaacgcccagaagcccagaacagacagataca 4560
DB 4501 ctgttaaccttgccttaataaagaacgcccagaagcccagaacagacagataca 4560
QY 4561 cccctagaagactgycgaagagataaagaagatagacagttctctgagactccggaggga 4620
DB 4561 cccctagaagactgycgaagagataaagaagatagacagttctctgagactccggaggga 4620
QY 4621 cctgctataccctacatgtggaggaatctctgcccacaagaaggttagaatagtcc 4680
DB 4621 cctgctataccctacatgtggaggaatctctgcccacaagaaggttagaatagtcc 4680
QY 4681 aacagatcatgcttcaacccacactaggaactaaacactcaagagcttgcagaatat 4740
DB 4681 aacagatcatgcttcaacccacactaggaactaaacactcaagagcttgcagaatat 4740
QY 4741 cccctatcatgttctgaggtacagagagtgctgacatggtgttcaaaactgtgtgc 4800
DB 4741 cccctatcatgttctgaggtacagagagtgctgacatggtgttcaaaactgtgtgc 4800
QY 4801 cctgccaagctgttgaatgttaactcttccagaatacctccaggaagagactaagggaa 4860
DB 4801 cctgccaagctgttgaatgttaactcttccagaatacctccaggaagagactaagggaa 4860
QY 4861 gccacccaaggcgtcacctggagagtggaacttcaactggaagtaaaagccgctataacgaa 4920
DB 4861 gccacccaaggcgtcacctggagagtggaacttcaactggaagtaaaagccgctataacgaa 4920
QY 4921 acaaaatatactatgttctttagacacacttccaggttgggttagagagctataccta 4980
DB 4921 acaaaatatactatgttctttagacacacttccaggttgggttagagagctataccta 4980
QY 4981 agaaagagacttcaacggtgtgtgtaagaaatactgagagaaatttccaaagattg 5040

Dh 4981 agaaagagactcaacgltggtgctaaagaataactggaagaaattttlccaaagatttg 5040
Qy 5041 gaatacctaaggtaataaggtcagacaaatgtgtccagacttctgttgcacaggttaagttag 5100
Dh 5041 gaatccctaaggtaataaggtcagacaaatgtgtccagacttctgttgcacaggttaagttag 5100
Qy 5101 gactgcccagaataattggggaattgattggaactgtcatltgtcga taacagaccccaagct 5160
Dh 5101 gactgcccagaataattggggaattgattggaactgtcatltgtcga taacagaccccaagct 5160
Qy 5161 caggaacaggtagaagagatagatagaacacataaagaagaccccttaaccaaatltgacacag 5220
Dh 5161 caggaacaggtagaagagatagatagaacacataaagaagaccccttaaccaaatltgacacag 5220
Qy 5221 agactggtcattaaatgattgattggtctccctgccttggcttggcttttaagggttagagaca 5280
Dh 5221 agactggtcattaaatgattgattggtctccctgccttggcttggcttttaagggttagagaca 5280
Qy 5281 ccccttggaacagttctgggtctgaccccccataatgcatctacagggggaaccccccgtttg 5340
Dh 5281 ccccttggaacagttctgggtctgaccccccataatgcatctacagggggaaccccccgtttg 5340
Qy 5341 caggaattgtcctttgacacataagtgctgtatgtgtctgtcttccagaccttgttctctagtc 5400
Dh 5341 caggaattgtcctttgacacataagtgctgtatgtgtctgtcttccagaccttgttctctagtc 5400
Qy 5401 tcaagagcgtctcgaatgaggtgagcagcagcgtggaagacagctcccggaagccctacacag 5460
Dh 5401 tcaagagcgtctcgaatgaggtgagcagcagcgtggaagacagctcccggaagccctacacag 5460
Qy 5461 gaaggaagcttgacaagtccacatctcgtcccaagtltgagaatltcaagtlcatgtatagaacgc 5520
Dh 5461 gaaggaagcttgacaagtccacatctcgtcccaagtltgagaatltcaagtlcatgtatagaacgc 5520
Qy 5521 accgttgagaagaacactcgaagactcgtggaagagacacttattctcgttaacttttgaacacac 5580
Dh 5521 accgttgagaagaacactcgaagactcgtggaagagacacttattctcgttaacttttgaacacac 5580
Qy 5581 caacggtctgttgaagaagtcgaaggaatcccccacactgcatccatgtcatccacglttaagccgg 5640
Dh 5581 caacggtctgttgaagaagtcgaaggaatcccccacactgcatccatgtcatccacglttaagccgg 5640
Qy 5641 cgcacactcccgactcgggtgtggaagcgaanaagactgagaatltcccttaagcttcgcgc 5700
Dh 5641 cgcacactcccgactcgggtgtggaagcgaanaagactgagaatltcccttaagcttcgcgc 5700
Qy 5701 tccctgcgtcgttcccttaacttaacataactcccccagcgcagtagtaaacgcttata 5760
Dh 5701 tccctgcgtcgttcccttaacttaacataactcccccagcgcagtagtaaacgcttata 5760
Qy 5761 gacacgtcgaacccccaataagacttatacccttaactgctggaattatltgaacccctgtagc 5820
Dh 5761 gacacgtcgaacccccaataagacttatacccttaactgctggaattatltgaacccctgtagc 5820
Qy 5821 ggtgtcaactgttaaatagactcgaaggtgtgtccttaagagacactggtgtgacctgaactg 5880
Dh 5821 ggtgtcaactgttaaatagactcgaaggtgtgtccttaagagacactggtgtgacctgaactg 5880
Qy 5881 catttcgtcctccgattgataaccccgctgttaaaagacacactcccaactcagtcgcgt 5940
Dh 5881 catttcgtcctccgattgataaccccgctgttaaaagacacactcccaactcagtcgcgt 5940
Qy 5941 agtataaggttctatattgtgtccagagacagagaagaagaataactcgttgggttcttg 6000
Dh 5941 agtataaggttctatattgtgtccagagacagagaagaagaataactcgttgggttcttg 6000
Qy 6001 gaatcctctgttagagatagagatcgtgtacactccaacagatggaacactggaatgtgcgg 6060
Dh 6001 gaatcctctgttagagatagagatcgtgtacactccaacagatggaacactggaatgtgcgg 6060
Qy 6061 atctctccagagccgggttaaaatctctctgttcaattccggcccggaagtacacaa 6120
Dh 6061 atctctccagagccgggttaaaatctctctgttcaattccggcccggaagtacacaa 6120

Qy 6121 atgtatgaactatataagaataagactgtctcccatcagacttaattatataagaata 6180
Dh 6121 atgtatgaactatataagaataagactgtctcccatcagacttaattatataagaata 6180
Qy 6181 agttcaactgaagaagaagaataataattcaaaagtgtatgaatgtatgctgtggg 6240
Dh 6181 agttcaactgaagaagaagaataataattcaaaagtgtatgaatgtatgctgtggg 6240
Qy 6241 aatagtatttlatatatagtcggtgggagcagaggttccacttlaacacatcgccttaagata 6300
Dh 6241 aatagtatttlatatatagtcggtgggagcagaggttccacttlaacacatcgccttaagata 6300
Qy 6301 gaagcggggaacagacccctctgtggaatgtggaacccgataaagttaactgtgtaacaggg 6360
Dh 6301 gaagcggggaacagacccctctgtggaatgtggaacccgataaagttaactgtgtaacaggg 6360
Qy 6361 ccccgcccttgtagcagcacacgataacttgcggttgcctcccaatlaacactcgcgtcgccct 6420
Dh 6361 ccccgcccttgtagcagcacacgataacttgcggttgcctcccaatlaacactcgcgtcgccct 6420
Qy 6421 gacataacacagcgccttagcacaacagtaaccactgtatltcccttaacacacgcttaga 6480
Dh 6421 gacataacacagcgccttagcacaacagtaaccactgtatltcccttaacacacgcttaga 6480
Qy 6481 aactcccgaggtgttctctgttlaagacagagaagacttccaagtctatccagaggact 6540
Dh 6481 aactcccgaggtgttctctgttlaagacagagaagacttccaagtctatccagaggact 6540
Qy 6541 ttccaaagcatcaactccaccccgacttgaatgtccacttctctgtgtgtctgtctatcc 6600
Dh 6541 ttccaaagcatcaactccaccccgacttgaatgtccacttctctgtgtgtctgtctatcc 6600
Qy 6601 tcaaggcccttataataagaaggtatggtcttaagaagaataatltcaatgtgaccaaagag 6660
Dh 6601 tcaaggcccttataataagaaggtatggtcttaagaagaataatltcaatgtgaccaaagag 6660
Qy 6661 catagaacatcaatgtatactggtgggttccggaataagacttaacctcaactgaagtltccggg 6720
Dh 6661 catagaacatcaatgtatactggtgggttccggaataagacttaacctcaactgaagtltccggg 6720
Qy 6721 aaggggacatgtataaggaagaagctcccccacatccacacacacttgcgtatagactgtg 6780
Dh 6721 aaggggacatgtataaggaagaagctcccccacatccacacacacttgcgtatagactgtg 6780
Qy 6781 gtttaagacaggtcctcagaanaatacagtatttaagtaactgtgtataacaggtgtgtggca 6840
Dh 6781 gtttaagacaggtcctcagaanaatacagtatttaagtaactgtgtataacaggtgtgtggca 6840
Qy 6841 tgaataactgggttlaaccccccgtgttltccacactcagttcaacacaaatccaaagatttc 6900
Dh 6841 tgaataactgggttlaaccccccgtgttltccacactcagttcaacacaaatccaaagatttc 6900
Qy 6901 tgtgtcatgtgtccaaatcgttcccccaggtgtataccaactcctgaggaagtgttcccttag 6960
Dh 6901 tgtgtcatgtgtccaaatcgttcccccaggtgtataccaactcctgaggaagtgttcccttag 6960
Qy 6961 gaataatgactcgttataacacgacacaaaagaagacccgtaacccttaaccttaactgtta 7020
Dh 6961 gaataatgactcgttataacacgacacaaaagaagacccgtaacccttaaccttaactgtta 7020
Qy 7021 atgtcgtatagtaggaacggtctgtggttaggaagaaggaacgctgacctgtagacagga 7080
Dh 7021 atgtcgtatagtaggaacggtctgtggttaggaagaaggaacgctgacctgtagacagga 7080
Qy 7081 ccaacagcagcttagagaagaagacttggtagactcaatltgcgcacatgtacagaagatctccga 7140
Dh 7081 ccaacagcagcttagagaagaagacttggtagactcaatltgcgcacatgtacagaagatctccga 7140
Qy 7141 gacttaagaaggtctgttagcaacactagaagaagttcccgacttcttgtctgaaggtgtt 7200
Dh 7141 gacttaagaaggtctgttagcaacactagaagaagttcccgacttcttgtctgaaggtgtt 7200

QY 193 gacttggaggtaaagagatcccccttggaaacgtgtgtgttcgcccgcgtctctgt 252
|||||
Db 181 gacttggaggtaaagagatcccccttggaaacgtgtgtgttcgcccgcgtctctgt 240
QY 253 ctgagatgtctgttttcgttgtatgagcgttcctcggtttgaaagtctctcaagacgttaa 312
|||||
Db 241 ctgagatgtctgttttcgttgtatgagcgttcctcggtttgaaagtctctcaagacgttaa 300
QY 313 ggaactggagagacttgcatacagaacagcttgcagagagatccaaagcttgcaccccttgggg 372
|||||
Db 301 ggaactggagagacttgcatacagaacagcttgcagagagatccaaagcttgcaccccttgggg 360
QY 373 acgcccggagagtgaggagagacagagagacgcttggtgtctccatagctgtgtcagag 432
|||||
Db 361 acgcccggagagtgaggagagacagagagacgcttgggtgtctccatagctgtgtcagag 420
QY 433 accgaatctctgtttgttgaagacgaaagctccccctccggcggtccgaactctttgct 492
|||||
Db 421 accgaatctctgtttgttgaagacgaaagctccccctccggcggtccgaactctttgct 480
QY 493 gcttggaaagacgagacgagctgcgctgtgtctgcagatctgtgttctgtctctgtgtg 552
|||||
Db 481 gcttggaaagacgagacgagctgcgctgtgtgtctgcagatctgtgttctgtctctgtgtg 540
QY 553 tcttctgtctgttgcgtctctgttctacagttttaataatgagacagatgactaaccgcc 612
|||||
Db 541 tcttctgtctgttgcgtctctgttctacagttttaataatgagacagatgactaaccgcc 600
QY 613 ttaatttgactctgcacatctgagcttgaagttaataccaaagctcacaatttgcagttc 672
|||||
Db 601 ttaatttgactctgcacatctgagcttgaagttaataccaaagctcacaatttgcagttc 660
QY 673 aggttaagaaagagaccttgcagacatctctgtcctctgaaatgagccaaacttgcattg 732
|||||
Db 661 aggttaagaaagagacaccttgcagacatctctgtcctctgaaatgagccaaacttgcattg 720
QY 733 gatgtgcatacagagagggacacttaattctgaaattatctctgtgcgtgttaaggaatactt 792
|||||
Db 721 gatgtgcatacagagagggacacttaattctgaaattatctctgtgcgtgttaaggaatactt 780
QY 793 ttcaagactggagacccggcgtctcatctctatcagagaccccttatcttaagtgtgcagatc 852
|||||
Db 781 ttcaagactggagacccggcgtctcatctctatcagagaccccttatcttaagtgtgcagatc 840
QY 853 tggcagaagaatcctccgcagatgggtttaacacatggtcctaaataaaccagaagaagccagatc 912
|||||
Db 841 tggcagaagaatcctccgcagatgggtttaacacatggtcctaaataaaccagaagaagccagatc 900
QY 913 ccggaatccctgtgctcttggagagaaaaaacaacatctgcgcgaaaaaagtcgagccctctt 972
|||||
Db 901 ccggaatccctgtgctcttggagagaaaaaacaacatctgcgcgaaaaaagtcgagccctctt 960
QY 973 cctcgtatctaccccgagatcgagagacgcccgaactgtgcgcgaaaccccaactgttctcc 1032
|||||
Db 961 cctcgtatctaccccgagatcgagagacgcccgaactgtgcgcgaaaccccaactgttctcc 1020
QY 1033 ccaccccttatctcaagacagagtgctgtgtgagggagacacttgcctccctcgagactccgg 1092
|||||
Db 1021 ccaccccttatctcaagacagagtgctgtgtgagggagacacttgcctccctcgagactccgg 1080
QY 1093 tgggtggagggagacctgtcgtccggagactcggagccggagagcgccaccccgagcggagcag 1152
|||||
Db 1081 tgggtggagggagacctgtcgtccggagactcggagccggagagcgccaccccgagcggagcag 1140
QY 1153 acgagatcgcgagatatccgctgagacatgagccctcccatgccaagggggccaatttgc 1212
|||||
Db 1141 acgagatcgcgagatatccgctgagacatgagccctcccatgccaagggggccaatttgc 1200
QY 1213 agccctcccgatattggcccttctctctgcagatcctataaattggaaaaactaacacatc 1272
|||||
Db 1201 agccctcccgatattggcccttctctctgcagatcctataaattggaaaaactaacacatc 1260
QY 1273 ccccttctcggagagatccccaacgctcgaagggtgtgtgagatccctatagttctctctc 1332

Db 1261 ccccttctcggagagatccccaacgctcgaagggtgtgtgagatccctatagttctctctc 1320
QY 1333 accagctacttggatgatatgtcaacagctgtctgcagacactcttcaacaacccagagagc 1392
|||||
Db 1321 accagctacttggatgatatgtcaacagctgtctgcagacactcttcaacaacccagagagc 1380
QY 1393 gagagagaatctctgtttagagagcttagaaaaatgttccctgggcccgaacggcgacccagc 1452
|||||
Db 1381 gagagagaatctctgtttagagagcttagaaaaatgttccctgggcccgaacggcgacccagc 1440
QY 1453 agttggcaaaatggagatggacatgttcccttgcactcccccgttttggactataaaca 1512
|||||
Db 1441 agttggcaaaatggagatggacatgttcccttgcactcccccgttttggactataaaca 1500
QY 1513 cggctgaagtgaggagagcttgaanaatctatccgaagctctgtgtgcggtctccggg 1572
|||||
Db 1501 cggctgaagtgaggagagcttgaanaatctatccgaagctctgtgtgcggtctccggg 1560
QY 1573 ggcgcctcaagacggccacatatttggctaaagttaagagagtgatgcagggagccgaacg 1632
|||||
Db 1561 ggcgcctcaagacggccacatatttggctaaagttaagagagtgatgcagggagccgaacg 1620
QY 1633 aacctccctcgatattcttgaagagctcaatggaaaccccttcaagcggatccaaccccttgg 1692
|||||
Db 1621 aacctccctcgatattcttgaagagctcaatggaaaccccttcaagcggatccaaccccttgg 1680
QY 1693 atccctacatagagccccaagaaagcctcagttgcctgtgcctcatattgggcagtcgcctc 1752
|||||
Db 1681 atccctacatagagccccaagaaagcctcagttgcctgtgcctcatattgggcagtcgcctc 1740
QY 1753 tggatatcaagaaagaaactcagaagcttgaagagttcaagagagcttgagttacgtgac 1812
|||||
Db 1741 tggatatcaagaaagaaactcagaagcttgaagagttcaagagagcttgagttacgtgac 1800.
QY 1813 taagtgaagagagagagaggtgttatacagaagggagacagaagagagagagagagagag 1872
|||||
Db 1801 taagtgaagagagagagaggtgttatacagaagggagacagaagagagagagagagagag 1860
QY 1873 gaaaaaagaaagat 1932
|||||
Db 1861 gaaaaaagaaagat 1920
QY 1933 tgaactaagatcttgcgcgcagctgtgtgaagggagagagagagagagagagagagagat 1992
|||||
Db 1921 tgaactaagatcttgcgcgcagctgtgtgaagggagagagagagagagagagagagagat 1980
QY 1993 tttagaaaattagtlcagagcccttagacagtlcaggggaaccttggcacaataagaccacatcg 2052
|||||
Db 1981 tttagaaaattagtlcagagcccttagacagtlcaggggaaccttggcacaataagaccacatcg 2040
QY 2053 acaagagacaggtgtgtgtattgttaagaaaaagacacatcgggcagaaggaactgcaccaaga 2112
|||||
Db 2041 acaagagacaggtgtgtgtattgttaagaaaaagacacatcgggcagaaggaactgcaccaaga 2100
QY 2113 agggaaacaaagagacgaaaggtccttagctctagaagaaataaagattagggagagagcgg 2172
|||||
Db 2101 agggaaacaaagagacgaaaggtccttagctctagaagaaataaagattagggagagagcgg 2158
QY 2173 gttcggacccctcccgagcccgaggtlaactttgaaggttggaggggcaaacagttgagt 2232
|||||
Db 2159 gttcggacccctcccgagcccgaggtlaactttgaaggttggaggggcaaacagttgagt 2218
QY 2233 tctcgtttgataccagagagggagatctcagtgctctacaacatlaggaaaaactaaag 2292
|||||
Db 2219 tctcgtttgataccagagagggagatctcagtgctctacaacatlaggaaaaactaaag 2278
QY 2293 aaaaaaaatccctgggttgaaggtgtccaaagggcaacggcagatataccatggactaccgaa 2352
|||||
Db 2279 aaaaaaaatccctgggttgaaggtgtccaaagggcaacggcagatataccatggactaccgaa 2338
QY 2353 gaacggttgaacttggagagtgagacggtaaaccaatcttctcgttcaatccctgagtgcc 2412
|||||

Db 2339 gaacctgtgacttgaggatggagcggglaaacccactcgtcttctgtgcatcccttgatgctc 2398
QY 2413 cagtaaccctctaggttagagactactgacccaagatgaggagctccaattctcttgaaac 2472
Db 2339 cagtaaccctctaggttagagactactgacccaagatgaggagctccaattctcttgaaac 2458
QY 2473 aaggaagacccaagagctgctgtgtgaatacaaacccactactgtgtgtgacccctccaat 2532
Db 2459 aaggaagacccaagagctgctgtgtgaatacaaacccactactgtgtgtgacccctccaat 2518
QY 2533 atgattgatactgactatattctccccaagttaaagcctgtgtaagaatatacagtcctgct 2592
Db 2519 atgattgatactgactatattctccccaagttaaagcctgtgtaagaatatacagtcctgct 2578
QY 2593 tggagcagcttccccaagcctgaggcagaacccgacaggaatgaggttttgccaagaactc 2652
Db 2579 tggagcagcttccccaagcctgaggcagaacccgacaggaatgaggttttgccaagaactc 2638
QY 2653 ccccaacaggttaattcaactgaaagcagtgctacacacagatatacagtcagtaaccct 2712
Db 2639 ccccaacaggttaattcaactgaaagcagtgctacacacagatatacagtcagtaaccct 2698
QY 2713 tgaatagaagagctcgaagaaggaatttggcgcagatgttcaaaagattaatccacaagggca 2772
Db 2699 tgaatagaagagctcgaagaaggaatttggcgcagatgttcaaaagattaatccacaagggca 2758
QY 2773 tccatgcttccctgccaatccctctggaatactccctctgctacacggttaggaagaactggga 2832
Db 2759 tccatgcttccctgccaatccctctggaatactccctctgctacacggttaggaagaactggga 2818
QY 2833 ccaatgattatcgacccaagatatacaggaacttgagaagagtcacaataaagggtgcagagacac 2892
Db 2819 ccaatgattatcgacccaagatatacaggaacttgagaagagtcacaataaagggtgcagagacac 2878
QY 2893 acccaacagctccgaaccccttatatacctcttgagcgccctcccgccggaacggaacagctgct 2952
Db 2879 acccaacagctccgaaccccttatatacctcttgagcgccctcccgccggaacggaacagctgct 2938
QY 2953 acacagatattgagacttaaaagatgctctctctgctgagattacacccactagaccac 3012
Db 2939 acacagatattgagacttaaaagatgctctctctgctgagattacacccactagaccac 2998
QY 3013 cacttttgccttgcgaatgagagatatacaggttaggaagaacccgggacactacactggga 3072
Db 2999 cacttttgccttgcgaatgagagatatacaggttaggaagaacccgggacactacactggga 3058
QY 3073 cccgacgtccccaaggggttcaagaactcccgacacatcttctgaacgaacccctacacaggg 3132
Db 3059 cccgacgtccccaaggggttcaagaactcccgacacatcttctgaacgaacccctacacaggg 3118
QY 3133 accgtgccaacttagagatccaacacccctcaggtgagacccctcccgtaagtgatgagac 3192
Db 3119 accgtgccaacttagagatccaacacccctcaggtgagacccctcccgtaagtgatgagac 3178
QY 3193 tgcctctggcgaggacccaacaaacagagctgcttagaaggtaacgaagcactactgctgg 3252
Db 3179 tgcctctggcgaggacccaacaaacagagctgcttagaaggtaacgaagcactactgctgg 3238
QY 3253 aattgtctgacactaggtacagagcctctgtctaagaagccacagatttgcagagagaggg 3312
Db 3239 aattgtctgacactaggtacagagcctctgtctaagaagccacagatttgcagagagaggg 3298
QY 3313 taacatacttggtgtacagtttggcgggcgaggcagatgctgacgggaagcaggaaga 3372
Db 3299 taacatacttggtgtacagtttggcgggcgaggcagatgctgacgggaagcaggaaga 3358
QY 3373 aaactgtagtcacaataccggcccaacacacagccaacaaagtgagagagatttttgggga 3432
Db 3359 aaactgtagtcacaataccggcccaacacacagccaacaaagtgagagagatttttgggga 3418
QY 3433 cagctgagatttgcagagctgtgataccgggggttgcgacacttagcagcccaactaac 3492
Db 3419 cagctgagatttgcagagctgtgataccgggggttgcgacacttagcagcccaactaac 3478

QY 3493 cgttaaccaaagaaaggggagattccctgggctccttgagacccaagagcatttgat 3552
Db 3479 cgttaaccaaagaaaggggagattccctgggctccttgagacccaagagcatttgat 3538
QY 3553 ctatacaaaaggcctgtctgagcgacactgtctgtgcccctcctgacgttaactaaacct 3612
Db 3539 ctatacaaaaggcctgtctgagcgacactgtctgtgcccctcctgacgttaactaaacct 3598
QY 3613 ttaccccttatgtgagtgagcgtaaggagtgagcccggaagagttttaacccaacactag 3672
Db 3599 ttaccccttatgtgagtgagcgtaaggagtgagcccggaagagttttaacccaacactag 3658
QY 3673 gacacatgagagagacctgtctgacactgtgtcaaaagcattgataccctgtaagcagtggt 3732
Db 3659 gacacatgagagagacctgtctgacactgtgtcaaaagcattgataccctgtaagcagtggt 3718
QY 3733 ggcctgtatgtctgaaagctatcgacagctgtgtgacatactgtgtcaagagcctgacaat 3792
Db 3719 ggcctgtatgtctgaaagctatcgacagctgtgtgacatactgtgtcaagagcctgacaat 3778
QY 3793 tgactttggacagaataataactgtlaatatgcccccaatgcatgtgagacaatcgttcggc 3852
Db 3779 tgactttggacagaataataactgtlaatatgcccccaatgcatgtgagacaatcgttcggc 3838
QY 3853 agccccaagacagatgagatgacccaacgcccgaatgacccaatacaaaagcctgtctca 3912
Db 3839 agccccaagacagatgagatgacccaacgcccgaatgacccaatacaaaagcctgtctca 3898
QY 3913 cagaaggggtacatttgcctccacagccgctctcaaacctgtgacacttcttcgcttgag 3972
Db 3899 cagaaggggtacatttgcctccacagccgctctcaaacctgtgacacttcttcgcttgag 3958
QY 3973 agactgatgaacacagtgactatgatatgccaatactatgattgagagagactggggcc 4032
Db 3959 agactgatgaacacagtgactatgatatgccaatactatgattgagagagactggggcc 4018
QY 4033 gcaaggaacttaacagacataccgctgacttgagaagtgctaaactgttctactgacggaa 4092
Db 4019 gcaaggaacttaacagacataccgctgacttgagaagtgctaaactgttctactgacggaa 4078
QY 4093 gcaaggaacttggttggaagtgagaagatgagctggggcgagtggtggaacggaccca 4152
Db 4079 gcaaggaacttggttggaagtgagaagatgagctggggcgagtggtggaacggaccca 4138
QY 4153 cgatctgggccaagcagctgtccgggaaggaacttcaagcgcaaaaggctgagctcagggcc 4212
Db 4139 cgatctgggccaagcagctgtccgggaaggaacttcaagcgcaaaaggctgagctcagggcc 4198
QY 4213 tcaaggaacttgcggtgcggtcggaagggaattcataaacttatacgaacagaggt 4272
Db 4199 tcaaggaacttgcggtgcggtcggaagggaattcataaacttatacgaacagaggt 4258
QY 4273 atgcttctgacactgacacagtaacagcgggccatctataaanaaagggtgtgctaacct 4332
Db 4259 atgcttctgacactgacacagtaacagcgggccatctataaanaaagggtgtgctaacct 4318
QY 4333 cagcagggaggggaataaagaacaaagggaattctaagctatataaagccttaact 4392
Db 4319 cagcagggaggggaataaagaacaaagggaattctaagctatataaagccttaact 4378
QY 4393 tgcacaaagagcttagctataatacactgtcccggaacatcaagaagccaaagatccat 4452
Db 4379 tgcacaaagagcttagctataatacactgtcccggaacatcaagaagccaaagatccat 4438
QY 4453 cttagaggaacacagatgagctgacaggtgtgcaaacagagcagccagcgtttaaacttc 4512
Db 4439 cttagaggaacacagatgagctgacaggtgtgcaaacagagcagccagcgtttaaacttc 4498
QY 4513 tgcctataataaagcccaagcccaagcccaagacccagacgagatgacacccctagaagact 4572
Db 4499 tgcctataataaagcccaagcccaagcccaagacccagacgagatgacacccctagaagact 4558

QY 4573 ggcagagataaaagataagaccagcttctctgaagctcccgagggagacctgtataact 4632
|||||
Db 4559 ggcagagataaaagataagaccagcttctctgaagctcccgagggagacctgtataact 4618
QY 4633 catactggaggaatccctctgcccacaaagaaggttagaatactgtccaaagatacatc 4692
|||||
Db 4619 catactggaggaatccctctgcccacaaagaaggttagaatactgtccaaagatacatc 4678
QY 4693 gtccaaaccacactaggaactaaacactcgacagctgtgtccaaacataccctatacatg 4752
|||||
Db 4679 gtccaaaccacactaggaactaaacactcgacagctgtgtccaaacataccctatacatg 4738
QY 4753 ttctgaagctacacagagagtgctgaactcggtgtgtccaaactgtgtcccttcgacgtcg 4812
|||||
Db 4739 ttctgaagctacacagagagtgctgaactcggtgtgtccaaactgtgtcccttcgacgtcg 4798
QY 4813 ttaatgtctaactcctccagaataactcccaagaaagagactaaagggagagccacccagcg 4872
|||||
Db 4799 ttaatgtctaactcctccagaataactcccaagaaagagactaaagggagagccacccagcg 4858
QY 4873 ctccactggagagtgacttaacttaagtaagacgggttaatacggaaacaaatatactat 4932
|||||
Db 4859 ctccactggagagtgacttaacttaagtaagacgggttaatacggaaacaaatatactat 4918
QY 4933 tggctttgtagacactcttccagatcggttagagagcttactactaaagaagagact 4992
|||||
Db 4919 tggctttgtagacactcttccagatcggttagagagcttactactaaagaagagact 4978
QY 4993 caaccgtgtgtgctaagaaatactctgagagaaattttccagaatttggaataactaaag 5052
|||||
Db 4979 caaccgtgtgtgctaagaaatactctgagagaaattttccagaatttggaataactaaag 5038
QY 5053 taataaggtacagaaatgtgtccagcttctgtgcccaggttaagtaagggagcttgccaga 5112
|||||
Db 5039 taataaggtacagaaatgtgtccagcttctgtgcccaggttaagtaagggagcttgccaga 5098
QY 5113 tatgggagattgatttggaactgtgactgtgacatacagaccocaaagcttcggagcagtag 5172
|||||
Db 5099 tatgggagattgatttggaactgtgactgtgacatacagaccocaaagcttcggagcagtag 5158
QY 5173 agagagatgaatagaaacaaatlaaagaagaccttaccaaattgacacacagagactgycatla 5232
|||||
Db 5159 agagagatgaatagaaacaaatlaaagaagaccttaccaaattgacacacagagactgycatla 5218
QY 5233 atgatttgatgagctctctctgccccttctgtctttagaggtgaagaaacaccttggaagct 5292
|||||
Db 5219 atgatttgatgagctctctctgccccttctgtctttagaggtgaagaaacaccttggaagct 5278
QY 5293 ttggagctaacccctatgaattgtctctacagggagaccccccgtctgacgaataattcct 5352
|||||
Db 5279 ttggagctaacccctatgaattgtctctacagggagaccccccgtctgacgaataattcct 5338
QY 5353 ttgcaacatgctgctgctgtctgtcttccagccttctctctataggctcaagcgctcg 5412
|||||
Db 5339 ttgcaacatgctgctgctgtctgtcttccagccttctctctataggctcaagcgctcg 5398
QY 5413 agtgggtgagagcagcgagcggtgaagacgctccggagggagctactccggggagagcttcg 5472
|||||
Db 5399 agtgggtgagagcagcgagcggtgaagacgctccggagggagctactccggggagagcttcg 5458
QY 5473 aagttccacatcgcttccaaagtctgagatcagctcctatgttagagcgcacacgtlcaagaa 5532
|||||
Db 5459 aagttccacatcgcttccaaagtctgagatcagctcctatgttagagcgcacacgtlcaagaa 5518
QY 5533 aactcgaactcggltggaagagaccttactcgtactttgacacacaaacggtlga 5592
|||||
Db 5519 aactcgaactcggltggaagagaccttactcgtactttgacacacaaacggtlga 5578
QY 5593 aagtcgaaggaatcccaacccggtatcagatcaccaggttaagcgggagccacactccg 5652
|||||
Db 5579 aagtcgaaggaatcccaacccggtatcagatcaccaggttaagcgggagccacactccg 5638
QY 5653 actcggggtgagagccgaaagactgaatactcccttaagcttcgctccatcgctcg 5712
|||||
Db 5639 actcggggtgagagccgaaagactgaatactcccttaagcttcgctccatcgctcg 5698
QY 5713 ttccctactctaacaataactccccaagcgagtagtaaacgcttataagacgctcgagac 5772
|||||
Db 5699 ttccctactctaacaataactccccaagcgagtagtaaacgcttataagacgctcgagac 5758
QY 5773 ccccaatagaccttataccttactccctgagctgatatctgaacccgatacgggtgtactgta 5832
|||||
Db 5759 ccccaatagaccttataccttactccctgagctgatatctgaacccgatacgggtgtactgta 5818
QY 5833 aatagacactcgaggtgtgtctctctagagagacactgtgtgctgaaactgacttctgctc 5892
|||||
Db 5819 aatagacactcgaggtgtgtctctctagagagacactgtgtgctgaaactgacttctgctc 5878
QY 5893 cgattgataaaccccgctgttlaaagcacactcccaactagtcogtaagttagtgc 5952
|||||
Db 5879 cgattgataaaccccgctgttlaaagcacactcccaactagtcogtaagttagtgc 5938
QY 5953 tatgtctgcccagagcagagaaagataactgtgggggtctctggggaatccctctgt 6012
|||||
Db 5939 tatgtctgcccagagcagagaaagataactgtgggggtctctggggaatccctctgt 5998
QY 6013 agagagatgagctgtctcaactccaaagatgagagactggaatggcgcatactctccag 6072
|||||
Db 5999 agagagatgagctgtctcaactccaaagatgagagactggaatggcgcatactctccag 6058
QY 6073 gaccgggtataatcttcccttctgtcaattccgcccggagaaatgaacaaatgtgaacta 6132
|||||
Db 6059 gaccgggtataatcttcccttctgtcaattccgcccggagaaatgtgaacta 6118
QY 6133 tataagataaagagctgtcccatcagacttagatactaaagataagttactgaa 6192
|||||
Db 6119 tataagataaagagctgtcccatcagacttagatactaaagataagttactgaa 6178
QY 6193 aggaataacaggaataatlatcaaaagltgaaatgtatgagctgagggaatagtttla 6252
|||||
Db 6179 aggaataacaggaataatlatcaaaagltgaaatgtatgagctgagggaatagtttla 6238
QY 6253 ttatatggcggggagacagaggttccactttaaaccattcgccttagaataagacgggga 6312
|||||
Db 6239 ttatatggcggggagacagaggttccactttaaaccattcgccttagaataagacgggga 6298
QY 6313 gaacccctgttgcaatggagaccgaataagtaactgtcgaaacagagggcccgccctg 6372
|||||
Db 6299 gaacccctgttgcaatggagaccgaataagtaactgtcgaaacagagggcccgccctg 6358
QY 6373 gagccacgcatactgtccggtgcccgaatlaacctcgtcgccgtgacataacag 6432
|||||
Db 6359 gagccacgcatactgtccggtgcccgaatlaacctcgtcgccgtgacataacag 6418
QY 6433 ccgccttagacacagtaacactgattgtatcttccaaacacacgcttagaactcccaagct 6492
|||||
Db 6419 ccgccttagacacagtaacactgattgtatcttccaaacacacgcttagaactcccaagct 6478
QY 6493 gtccctgttaagacagagacagactcctcagctcatccagggagcttccaaagcactc 6552
|||||
Db 6479 gtccctgttaagacagagacagactcctcagctcatccagggagcttccaaagcactc 6538
QY 6553 aactccacgacccctgactgacacttctctgtgtgtgtcttctacactcagggctcct 6612
|||||
Db 6539 aactccacgacccctgactgacacttctctgtgtgtgtgtcttctacactcagggctcct 6598
QY 6613 tattatgaggggagtggtctaagaaanaaaltcaatgtatccaaagagatagaatcaa 6672
|||||
Db 6599 tattatgaggggagtggtctaagaaanaaaltcaatgtatccaaagagatagaatcaa 6658
QY 6673 tgtacatgggggtcccgaaataagcttaccctcactgaagcttccgggaaggagacatgc 6732
|||||
Db 6659 tgtacatgggggtcccgaaataagcttaccctcactgaagcttccgggaaggagacatgc 6718
QY 6733 ataggaaaagctcccccactccacacacacttctgctatgactgtgttatagacag 6792
|||||

Db 6719 ataggaagaagctcccccatcccaacaacaccttgcctatagctggtttatgacag 6778
 QY 6793 gccctagaataacagattatgacctggtataacaggtggtgggcatgcaatactgg 6852
 Db 6779 gccctagaataacagattatgacctggtataacaggtggtgggcatgcaatactgg 6838
 QY 6853 ttaacccctgtgttccacctcagttttcaacccaatccaagaatttctgtgcatgctc 6912
 Db 6839 ttaacccctgtgttccacctcagttttcaacccaatccaagaatttctgtgcatgctc 6898
 QY 6913 caaatgctccccagagtgctactacatccctgaggaagtgttccctgagatatactat 6972
 Db 6899 caaatgctccccagagtgctactacatccctgaggaagtgttccctgagatatactat 6958
 QY 6973 cgggtataaccgacccaagaagaacccgtatccctaccctagctgtatgtctcgatata 7032
 Db 6959 cgggtataaccgacccaagaagaacccgtatccctaccctagctgtatgtctcgatata 7018
 QY 7033 gggagcggccgttggtggtggaacaggaagctgtccctgataccagagccacagatgta 7092
 Db 7019 gggagcggccgttggtggtggaacaggaagctgtccctgataccagagccacagatgta 7078
 QY 7093 gagaagaagagcttggtggtggaacaggaagctgtccctgataccagagccacagatgta 7152
 Db 7079 gagaagaagagcttggtggtggaacaggaagctgtccctgataccagagccacagatgta 7138
 QY 7153 tctgtttagcaacctagaagaagctccctgactcttctgttctgaagtgttctacagaacccg 7212
 Db 7139 tctgtttagcaacctagaagaagctccctgactcttctgttctgaagtgttctacagaacccg 7198
 QY 7213 aaggagattagatctgtcttcttaagaagaagtggttatgtgacgcttaagaagaaga 7272
 Db 7199 aaggagattagatctgtcttcttaagaagaagtggttatgtgacgcttaagaagaaga 7258
 QY 7273 tgtgtctctatgtagatcactcagagagccatcagagatccatgagaagaagcttagaaga 7332
 Db 7259 tgtgtctctatgtagatcactcagagagccatcagagatccatgagaagaagcttagaaga 7318
 QY 7333 aagtagagagagctgctcgaagaagaagaagctgacccgggtgtgttgaagaagtgtctc 7392
 Db 7319 aagtagagagagctgctcgaagaagaagaagctgacccgggtgtgttgaagaagtgtctc 7378
 QY 7393 aacagctcctctgtgatacagaccctgtcttctgtcgtcgaaggggccccctagtagtccgtg 7452
 Db 7379 aacagctcctctgtgatacagaccctgtcttctgtcgtcgaaggggccccctagtagtccgtg 7438
 QY 7453 ctctgttacttacaggttggtggtctgtcttaataagttgtgtccttctgttagagaa 7512
 Db 7439 ctctgttacttacaggttggtggtctgtcttaataagttgtgtccttctgttagagaa 7498
 QY 7513 cgaagtgagtgagtgccagatcatatgttacttaggcaacagtaaccagagcctctgagccaa 7572
 Db 7499 cgaagtgagtgagtgccagatcatatgttacttaggcaacagtaaccagagcctctgagccaa 7558
 QY 7573 ggaagaactgagcctctagccttcccaagcttctaagattagaactttaaagaagaagaag 7632
 Db 7559 ggaagaactgagcctctagccttcccaagcttctaagattagaactttaaagaagaagaag 7618
 QY 7633 tggggaatgaaagatgaaatgcaacctaaacctcccaagaaccaggaagttaataaa 7692
 Db 7619 tggggaatgaaagatgaaatgcaacctaaacctcccaagaaccaggaagttaataaa 7678
 QY 7693 agctctaagtccccgaatactaacagaccgtgctgctgcaagtaaatagtagaaggtca 7752
 Db 7679 agctctaagtccccgaatactaacagaccgtgctgctgcaagtaaatagtagaaggtca 7738
 QY 7753 caactctattgttccagggcctctatccctggcctaagaagaatacagaagaatggtt 7812
 Db 7739 caactctattgttccagggcctctatccctggcctaagaagaatacagaagaatggtt 7798
 QY 7813 gactaatcgtctatctgtagtctgtlaaaactgtagcacacatagaagaattgattacac 7872
 Db 7799 gactaatcgtctatctgtagtctgtlaaaactgtagcacacatagaagaattgattacac 7858

QY 7873 atgacagcccttagtgacctatctcaactgcaatctgttcaactctgcccagaagccacgc 7932
 Db 7859 atgacagcccttagtgacctatctcaactgcaatctgttcaactctgcccagaagccacgc 7918
 QY 7933 agatgagcactccgagagctattttaaagtatgtgtccacgagcggtgctctcgata 7992
 Db 7919 agatgagcactccgagagctattttaaagtatgtgtccacgagcggtgctctcgata 7978
 QY 7993 tttaaaatgatgtgtccatggaagcggtgctctcgatattttaaagtatgtgtgtg 8052
 Db 7979 tttaaaatgatgtgtccatggaagcggtgctctcgatattttaaagtatgtgtgtg 8038
 QY 8053 acgcacagcgttctgttgtaacccataaagctgttcccgatctcgacccgagcgccga 8112
 Db 8039 acgcacagcgttctgttgtaacccataaagctgttcccgatctcgacccgagcgccga 8098
 QY 8113 gtctctaccccgctgtgtgtgtagactgtgtggtccacgagcggttgaataaatcct 8172
 Db 8099 gtctctaccccgctgtgtgtgtagactgtgtggtccacgagcggttgaataaatcct 8158
 QY 8173 ctgtgcttgcattcaaaaaaataaaaaaataaaaaa 8209
 Db 8159 ctgtgcttgcattcaaaaaaataaaaaaataaaaaa 8195
 RESULT 3
 AAF77727
 ID AAF77727 standard; DNA; 8132 BP.
 XX
 AC AAF77727;
 XX
 DT 23-MAY-2001 (first entry)
 XX
 DE Nucleotide sequence of a retrovirus found in miniature swine.
 XX
 KW Retrovirus; graft transplantation; xenotransplantation; miniature swine;
 XX
 OS Unidentified.
 XX
 PN US6190861-B1.
 XX
 PD 20-FEB-2001.
 XX
 PF 13-DEC-1996; 9605-0766528.
 XX
 PR 14-DEC-1995; 95US-0572645.
 XX
 PA (GENO) GEN HOSPITAL CORP.
 XX
 PI Fishman JA;
 XX
 DR WPI: 2001-256211/26.
 XX
 PT P-PSDB; AAB73285, AAB73286, AAB73287.
 XX
 PT Assessing risk of endogenous retroviruses in clinical practice and in
 XX xenotransplantation, comprises using probe sequences derived from swine
 XX or miniature swine retroviral genome -
 PS Claim 1; Fig 3; 127pp. English.
 CC The present invention relates to a method for screening a cell or tissue
 CC for the presence or expression of a retrovirus (RV), comprising
 CC contacting a target nucleic acid from the cell or tissue with a second
 CC nucleic acid from the present invention (e.g. the present sequence or a
 CC fragment thereof). The method is useful for RV detection and to assess
 CC graft transplantation risk. Screening of animals allows the elimination
 CC of donors with active replication of known viruses. Inactive proviruses
 CC can be detected and inactivated, allowing identification and elimination
 CC of potential human pathogens derived from swine in a manner not possible
 CC in the outbred human organ donor population and is important to the
 CC development of human xenotransplantation.

XX Sequence 8132 BP, 2248 A; 1977 C; 2037 G; 1870 T; 0 other;

Query Match 76.5%; Score 6282.4; DB 22; Length 8132;
Best Local Similarity 87.1%; Pred. No. 0;
Matches 7165; Conservative 0; Mismatches 931; Indels 134; Gaps 19;

QY 1 gttgtgtacgactgtggcccccagcgcttggaaataaaatccctctgtcttgcac 60
Db 3 gtgtgtacgactgtggcccccagcgcttggaaataaaatccctctgtcttgcac 62
QY 61 aagaccgcttctcgtgaatgatt-tgggtgtcgcctcttccagaccgaggggga 119
Db 63 aagaccgcttctcgtgaatgatttgaaggagtccttccagaccgaggg----- 115
QY 120 ttgtcttctacgtgccttcatattgtgcgttggccggaaatccctggaccctt 179
Db 116 ---ttcttctgtcgttctacatlttgggggtcgtccggatcgtccgcggccacctta 173
QY 180 aacccgagaaacgacttggaggtta-aaggatccctcttggaaacgtgtgtg--tgtgtc 236
Db 174 aacccgagaaacgacttggaggttaaaagaatccctctttaaagtgatgcgtgaacc 233
QY 237 ggcggggtctctgttctgaatgtctgttctcgttgatgcgccttcogtltgcagctg 296
Db 234 ggcggggtctctgttctgaatgtctgttctcgttgatgcgccttcogtltgcagctg 293
QY 297 tcccttcagacgtgaagactggagactgtgatcagagacggtctagaagagttcag 356
Db 294 tcccttcagacgtgaagactggagactgtgatcagagacggtctagaagagttcag 353
QY 357 gctggcaacgtggggagaccccgaggatgtggagagacccggagacgctgtgtgtctcc 416
Db 354 gctggcctcgtgggagaccccgaggatgtggagagacccggagacgctgtgtgtctcc 413
QY 417 tactgtcgtcagaaggagacggtctgttcttgaagcgaaagcttccctccgcggcgg 476
Db 414 tactgtcgtcagaaggagacggtctgttcttgaagcgaaagcttccctccgcggcgg 473
QY 477 tccgactcttctgcctgttcttggaaacgagcgggtccggtgtcttgatcttgg 536
Db 474 tccgactcttctgcctgttcttggaaacgagcgggtccggtgtcttgatcttgg 533
QY 537 gttctgtctcgtgtgtcttctgttcttggccttcttcaagtttaatatggagacg 596
Db 534 gttctgtctcgtgtgtcttctgttcttggccttcttcaagtttaatatggagacg 593
QY 597 acatgtactacccctctagtattgactctcgacatctgagactgaatlaacagagct 656
Db 594 acgtgtacgacccctcttaagtttgaactctcgacatctgagactgaatlaacagagct 653
QY 657 cataattgtcagttcaggttaagaaggacacttggacagacttcttgcctcgaatgg 716
Db 654 cataattgtcagttcaggttaagaaggacacttggacagacttcttgcctcgaatgg 713
QY 717 ccaacattcgaatgttggatggcaltcagaggagacctttaatcttgaaatlatcctggct 776
Db 714 ccgacattcgaatgttggatggcaltcagaggagacctttaatcttgaaatlatcctggct 773
QY 777 gtttaaggcaatcatttttcaagactggaccggcgtctctatccctgatcaggaacctatac 836
Db 774 gtttaaggcaatcatttttcaagactggaccggcgtctctatccctgatcaggaacctatac 833
QY 837 cttaacgtgcagaatttggcagaagatccctccgacatgggttaaaccaatgacctaaataa 896
Db 834 cttaacgtgcagaatttggcagaagatccctccgacatgggttaaaccaatgacctaaataa 893
QY 897 ccaagaaagcaggatcccgaaatccctgtgtcttggagagaaaaaacaacacatcggccgaa 956
Db 894 ccaagaaagcaggatcccgaaatccctgtgtcttggagagaaaaaacaacacatcggcgaa 953
QY 957 aaagtgcagccctcttccgtatctaaccccgagatcgagagacgctgcgcgaa 1016

Db 954 aaagtgcagccctc-tcctatatactaaccccgagatggagagccacggcttggccgaa 1012
QY 1017 accccaacactgttcccccacaccccttaccagaaaggtgtcgttgaaggga-cctctgc 1075
Db 1013 accccaacactgttcccccacaccccttaccagaaaggtgtcgttgaagggaaccttgc 1072
QY 1076 cctctcttgagactccggttggatggagacactgtctgcggagactcagaccgagagcg 1135
Db 1073 cctctcttgagactccggttggatggagacactgtctgcggagactcagaccgagagcg 1132
QY 1136 caacccgagcggagacagagatcgcgatatccgcgttgcgacactatgacctccat 1195
Db 1133 caacccgagcggagacagagatcgcgatatccgcgttgcgacactatgacctccat 1192
QY 1196 gccagggggcgaatttgaagccctccagatattgaccttctcttcgcagatctctata 1255
Db 1193 accggggggcgaatttgaagccctccagatattgaccttctcttcgcagatctctata 1252
QY 1256 ttggaataactaacatcccttctcggaggaatcccaacgacctacggggttggaa 1315
Db 1253 ttggaataactaacatcccttctcggaggaatcccaacgacctacggggttggaa 1312
QY 1316 gtccctatgttctctcacaagactacttggagatgtgtcaacagctgctcagacact 1375
Db 1313 gtccctatgttctctcacaagactacttggagatgtgtcaacagctgctcagacact 1372
QY 1376 ctccaacacggagagcgagagaaatctcgttaagggtctagaagaaatgcttctgggc 1435
Db 1373 ctccaacacggagagcgagagaaatctcgttaagggtctagaagaaatgcttctgggc 1432
QY 1436 cgaacggcgacacccacgacttgcgaataatgagatctgacatggatcttccctgtcgcgc 1495
Db 1433 cgaacggcgacacccacgacttgcgaataatgagatctgacatggatcttccctgtcgcgc 1492
QY 1496 cgggttggagactacacaacggtcgaaggatggagagcttgaataatctatccagagctct 1555
Db 1493 cgggttggagactacacaacggtcgaaggatggagagcttgaataatctatccagagctct 1552
QY 1556 ggttcggggtctcgcggggtccctcaagacgcccataatttggctcaaggtaaggaggt 1615
Db 1553 ggttcggggtctcgcggggtccctcaagacgcccataatttggctcaaggtaaggaggt 1612
QY 1616 gatgcagggagacgaacacactccctcgggtattcttcttggagagcctatgaaacgtctcag 1675
Db 1613 gatgcagggagacgaacacactccctcgggtattcttcttggagagcctatgaaacgtctcag 1672
QY 1676 gcggttcaaccccttctgacttcaactcagaggcccaagaagcctcagtggtccctggcctt 1735
Db 1673 gcggttcaaccccttctgacttcaactcagaggcccaagaagcctcagtggtccctggcctt 1732
QY 1736 catltygcagctgcgtcttgatatactgaaagaaacttcaagaaacttgaagaggttttcaagaa 1795
Db 1733 catltygcagctgcgtcttgatatactgaaagaaacttcaagaaacttgaagaggttttcaagaa 1792
QY 1796 ggtctggttactgtacttgaagagagcgaagaggtgttactaagaaggagagacaaga 1855
Db 1793 ggtctggttactgtacttgaagagagcgaagaggtgttactaagaaggagagacaaga 1852
QY 1856 agagggagagagaaacagagaaagaaagagagagagagaaagggagaaacgttgatag 1915
Db 1853 agagggagagagaaacagagaaagaaagagagagagagagaaagggagaaacgttgatag 1912
QY 1916 acggcagaagagaaatcttgaactaagaatcttggccgaggtgttgaagagaaagagagcag 1975
Db 1913 acggcagaagagaaatcttgaactaagaatcttggccgaggtgttgaagagaaagagagcag 1972
QY 1976 ggaagagagagagatlttaagaaatlaagtcagagcccttgaacgtcagggaaacctggg 2035
Db 1973 ggaagagagagagatlttaagaaatlaagtcagagcccttgaacgtcagggaaacctggg 2032
QY 2036 catatggacccacttcgaagaaagacaggtgtgtctatgttaagaaagaaagacactggc 2095

Db 2033 caataaggaccccaactcgacaaaggacaaatgctgacatctgttaagaagaagagacatgagc 2092
QY 2096 aaggactgcccccaagaagggaaaacaaaggacgaagcttactagctctagaagaataa 2155
Db 2093 aaggaactgcccccaagaagggaaaacaaaggacgaagcttactagctctagaagaataa 2152
QY 2156 agattaaaggagacgaggtctcgagccccctcccgagccagggtaactcttgaagtga 2215
Db 2153 agattaaaggagacgaggtctcgagccccctcccgagccagggtaactcttgaagtga 2212
QY 2216 ggggcaacccagttaagttccggttgaataccgagcggaacattcagttgtctacacac 2275
Db 2213 ggggcaacccagttaagttccggttgaataccgagcggaacattcagttgtctacacac 2272
QY 2276 attggaataaataaagaataaataatccgtggtatgtgtgacaaagcggaacgacgata 2335
Db 2273 attggaataaataaagaataaataatccgtggtatgtgtgacaaagcggaacgata 2332
QY 2336 tccatggaactacccgaagaacgcttgagctggagctggagcggttaacccactcgttct 2395
Db 2333 tccatggaactacccgaagaacgcttgagctggagctggagcggttaacccactcgttct 2392
QY 2396 ggtatccctgaagtcgcaagtaacccctctaggtagaagacttactgccaagatggagac 2455
Db 2393 ggtatccctgaagtcgcaagtaacccctctaggtagaagacttactgccaagatggagac 2452
QY 2456 tcaaatctcttgaacaaaggaaagacagagtgctcgtgaataacaaacccactcagtc 2515
Db 2453 acaaatctcttgaacaaaggaaagacagagtgctcgtgaataacaaacccactcagtc 2512
QY 2516 gttgaacctccaattagaatagaataatgcataatctccccaagtaaaagcttgata 2575
Db 2513 gttgaacctccaattagaatagaataatgcataatctccccaagtaaaagcttgata 2572
QY 2576 agatatacagtcggtgagagcttccccaagcctgggcaagaacccgagagatggg 2635
Db 2573 aatatacaatctcgtgttgagacagcttccccaagcctgggcaagaacccgagagatggg 2632
QY 2636 ttgggcaagaagcttccccaagctatctcaactgaagccagtgccacacagtgctc 2692
Db 2633 ttgggcaagaagcttccccaagctatctcaactgaagccagtgccacacagtgctc 2692
QY 2696 agtaagacagtaaccccttggtagaaggtctcgagaaggaattggcgcatgttcaag 2755
Db 2693 agtaagacagtaaccccttggtagaaggtctcgagaaggaattggcgcatgttcaag 2752
QY 2756 attaatccaacagagcactagcttccgttcccaatccctctggagatactccctgctac 2815
Db 2753 attaatccaacagagcactagcttccgttcccaatccctctggagatactccctgctac 2812
QY 2816 ggttaagaagccttggaacaaatgataatcgacacagtaacaggaacttgaagaggtcaataa 2875
Db 2813 ggttaagaagccttggaacaaatgataatcgacacagtaacaggaacttgaagaggtcaataa 2872
QY 2876 aagggttcagagacatacacaacaaaggtccggaacccctataaactcttgagcgccctcc 2935
Db 2873 aagggttcagagacatacacaacaaaggtccggaacccctataaactcttgagcgccctcc 2932
QY 2936 ggcgaacggaactggtacacagatgtggaacttaagaagtccttctcgccctgagat 2995
Db 2933 accccaacggaactggtacacagatgtggaacttaagaagtccttctcgccctgagat 2992
QY 2996 acaccaactagccaacactcttgccttcgaatgagagatccagtgacgggaagac 3052
Db 2993 acaccaactagccaacactcttgccttcgaatgagagatccagtgacgggaagac 3052
QY 3056 cgggcaactcaactggaacccgagctgccccgaagggtcaagaactcccgacactcttga 3115
Db 3053 cgggcaactcaactggaacccgagctgccccgaagggtcaagaactcccgacactcttga 3112
QY 3116 cgaagcctacacagagagctgccaacttcagatccaacacccctlaagttgacccctc 3175
Db 3113 cgaagcctacacagagagctgccaacttcagatccaacacccctlaagttgacccctc 3172

QY 3176 ccagtaactggaatgacactgcttctcgcggaagccacaacagagactcttgaagaagtac 3235
Db 3173 ccagtaactggaatgacactgcttctcgcggaagccacaacagagactcttgaagaagtac 3232
QY 3236 gaagcactactgcttgaatctgctgacatgatacagagctcttctgaagaagccca 3295
Db 3233 gaagcactactgcttgaatctgctgacatgatacagagctcttctgaagaagccca 3292
QY 3296 gatttcaggaagagagatacacttgggtgacagtttgcgggacggaacgata 3355
Db 3293 gatttcaggaagagagatacacttgggtgacagtttgcgggacggaacgata 3352
QY 3356 gacgagcgacaggaagaaactgtagtccagaataccgccccacaacagccaacagat 3415
Db 3353 gacgagcgacaggaagaaactgtagtccagaataccgccccacaacagccaacagat 3412
QY 3416 gaaagagcttcttgggagacagctggaatttgcagactgtgatacccggtttggaact 3475
Db 3413 gaaagagcttcttgggagacagctggaatttgcagactgtgatacccggtttggaact 3472
QY 3476 agcagcccaactctacccgtaacaaagaagaaggagatctctctgagctcgaaga 3535
Db 3473 agcagcccaactctacccgtaacaaagaagaaggagatctctctgagctcgaaga 3532
QY 3536 ccagaaggaatttgatgatacaaaaagccctgctgagcgacactgctcgccctcc 3595
Db 3533 ccagaaggaatttgatgatacaaaaagccctgctgagcgacactgctcgccctcc 3592
QY 3596 tgaagtaactaaaccccttacccttattgtgaataagcgttaaggagagacccggaagat 3655
Db 3593 tgaagtaactaaaccccttacccttattgtgaataagcgttaaggagagacccggaagat 3652
QY 3656 tttaaccacaacacttagagacataatgagagacacttgcctacactgtaaaaagactga 3715
Db 3653 tttaaccacaacacttagagacataatgagagacacttgcctacactgtaaaaagactga 3712
QY 3716 tccctgagccagtggttggcccggtatgtctgaagctatcgagctgtggccatactggt 3775
Db 3713 tccctgagccagtggttggcccggtatgtctgaagctatcgagctgtggccatactggt 3772
QY 3776 caaggaacgttgaacaaattgaactttggaagataatgaatgaagccccaatgaact 3835
Db 3773 caaggaacgttgaacaaattgaactttggaagataatgaatgaagccccaatgaact 3832
QY 3836 ggaagacactgcttgcgacagcccccagacagatgataatgacacagcccgatga 3895
Db 3833 ggaagacactgcttgcgacagcccccagacagatgataatgacacagcccgatga 3892
QY 3896 tcaagcctgcttctcaacagaagaaggtcaacttgcctcaacagcccgatga 3955
Db 3893 tcaagcctgcttctcaacagaagaaggtcaacttgcctcaacagcccgatga 3952
QY 3956 caactcttgccttgaagaagacagatgaacacagtgactaattgacatcaactatgat 4015
Db 3953 caactcttgccttgaagaagacagatgaacacagtgactaattgacatcaactatgat 4012
QY 4016 tgaagagacttgggttccggaagagacttcaacagacataacgctgagcttgaagaagtactaac 4075
Db 4013 tgaagagacttgggttccggaagagacttcaacagacataacgctgagcttgaagaagtactaac 4072
QY 4076 ctggttcaactgaacgaagcagctatgctgtgaaggtlaagagatgcttgggcgagct 4135
Db 4073 ctggttcaactgaacgaagcagctatgctgtgaaggtlaagagatgcttgggcgagct 4132
QY 4136 ggttgagcggaacccgagacatctgggcagagacgcttccggaaggaacttcaagcaaa 4195
Db 4133 ggttgagcggaacccgagacatctgggcagagacgcttccggaaggaacttcaagcaaa 4192
QY 4196 ggtcagctcatgagccctcaacgaagcttggcgcttggccgaaggaataccataaacat 4255
Db 4193 ggtcagctcatgagccctcaacgaagcttggcgcttggccgaaggaataccataaacat 4252

[illegible]

OY	7532	tcatggaacttaggcacacagctaccaggcccttcgtggccaaggaagaatcgacctcagc	7591
Db	7481	tcaatggaacttaggcacacagctaccaggcccttcgtggccaaggaagaatcgacctcagc	7533
OY	7592	cttcccagtttcaagtattgaactatataacaagaacagaagtgtgggaaatgtaaaagatga	7651
Db	7538	tctaaccggttctaagatctagaactatataacaagaagaagaagtgtgggaaatgtaaaagatga	7597
OY	7652	aatgcacataaacccctccagaacccccagaagttaataaaaagctctaagtcccccga	7711
Db	7598	aatacaaccttaagcttaatgtgaagcttaaatgttcttgtaactlccagaagtgtgttccta	7657
OY	7712	ttacagagccctcgctgctgcagtaaatagtgtgaagagtcacaccttccattgtttccag	7771
Db	7658	taggttaaaaagattaggtgttttttctgttlttaaataatgtcggaagttaaatagtccttag	7717
OY	7772	gccgtcataccctgycct-aagtaagataaacaggaatgaagtgtac---taatcgctat	7826
Db	7718	tacatgtctcttaggcataagaaactcttgtaactatttgtagataacagaagaagaggtt	7777
OY	7827	ctggaattctgtaaacactgactgtgcacataagaagaattgatatacattgtacagccctag	7886
Db	7778	ctaactgtctgtttaagcttctgttaaaaactgtgttcgcataaagaatgttgaagtgtgat	7837
OY	7887	tgacctatccaactgcatactgtcaactcgtgccaggaagcccaagatgtcgacccctc	7946
Db	7838	acaatatctctgtgtgaacaacatgtctcccccacccccaacatgtcgcaatgtgttaactc	7897
OY	7947	ggaagctattttaaattgattgtgtccacaggaacgctgcctcgcataattttaaattgattg	8006
Db	7898	taaacacatttaaatatttggtcccaagaagcgcgtcctcgaagttttaattgactgt	7957
OY	8007	gtcacatggagcgcgcgtcccgatattttaaattgattgtttgttgaacgcacaagcttgg	8066
Db	7958	gtttgt-----gatatttgaattgattgtttgttaaaagcgcgcgttgg	8002
OY	8067	ttgtgaaccccatlaaagctgtcccgatctccgaactcgggcgcgcagctccctacccctg	8126
Db	8003	ttgtgaaccccatlaaagctgtcccgatctccgaactcgggcgcgcagctccctacccctg	8063
OY	8127	cgtgtgttacgactgtgtggccccagcgcgcttggaaataaaalccctctgtctgtttgcat	8186
Db	8063	cgtgtgtgtacgactgtgtggccccagcgcgcttggaaataaaalccctctgtctgtttgcat	8122
OY	8187	caaaaaaaaaa	8196
Db	8123	caaaaaaaaa	8132
<hr/>			
RESULT 4			
AAT74884	ID	AAT74884 standard; cDNA; 7892 BP.	
XX	NC	AAT74884;	
XX	DT	09-FEB-1998 (first entry)	
XX	DE	Miniature swine retrovirus CDNA.	
XX	RW	Retrovirus; porcine; GAG protein; POL protein; ENV protein;	
KW	XO	xenotransplantation; infectious; provirus; organ transplant; donor;	
KW	activated virus; PCR; ss.		
XX	OS	Porcine retrovirus.	
XX	Key	Location/Qualifiers	
FT	CDS	585..2159	
FT		/+lag= a	
FT	mat_peptide	585..2156	
FT		/+lag= b	
FT		/note= "putative GAG protein"	
FT	CDS	2307..5744	

FT mat-peptide /tag= C
 2307..5741 /tag= d
 FT /note= "putative POL protein"
 FT CDS 5620..7536
 FT /tag= e
 FT mat-peptide 5620..7533
 FT /tag= f
 FT /note= "putative ENV protein"
 PN MO9721836-A1.
 PD 19-JUN-1997.
 PF 13-DEC-1996; 96WO-US19680.
 PR 14-DEC-1995; 95US-0572645.
 XX (GEHO) GEN HOSPITAL CORP.
 PA Fishman JA;
 PI MPI: 1997-332804/30.
 DR P-PSDB; AAM32096-W32098.
 XX
 XX New nucleic acid from porcine retroviruses - used for detecting
 PT viruses in transplant or other tissue and for assessing risk of
 PT transmitting infection to graft recipient
 XX
 PS Claim 22; Fig 3; 128pp; English.

CC This CDNA sequence represents a porcine retrovirus from miniature swine
 CC containing the putative coding regions for viral GAG, POL and ENV
 CC proteins. This sequence and PCR fragments generated from it
 CC (see AAT74812-T74882) can be used to screen organs for the presence of
 CC porcine retroviruses prior to xenotransplantation. Transplantation can
 CC increase the likelihood of retroviral activation if intact and
 CC infectious proviruses are present. The porcine retroviral sequence can be
 CC used to generate probes to determine the level (e.g. copy number) of
 CC intact (i.e. potentially replicating) porcine provirus sequences in a
 CC strain of xenograft transplantation donors. It can be used to detect
 CC mutations, genetic lesions or viral recombinants and also to determine
 CC the histological localisation of activated retroviruses. Using Polymerase
 CC Chain Reaction DNA Quantitation (PQ) on blood mononuclear cells,
 CC infectivity titration and susceptibility testing can be performed.
 CC Intimately animal donors without intact porcine retroviral sequences or a
 CC lower copy number of viral elements could be selected.

CC Sequence 7892 BP; 2191 A; 1915 C; 1980 G; 1806 T; 0 other;

Query Match 74.8%; Score 6141; DB 18; Length 7892;
 Best Local Similarity 89.0%; Pred. No. 0;
 Matches 6856; Conservative 0; Mismatches 735; Indels 114; Gaps 16;

QY 1 gtgtgtgtagcactgtggccagcgctgtggaataaaatcctctgtgtgtgac 60
 Db 3 gtgtgtgtagcactgtggccagcgctgtggaataaaatcctctgtgtgtgac 62
 QY 61 aagaccgttctgtgtgtgtgt-tgggtgtgtgtgtgtgtgtgtgtgtgtgt 119
 Db 63 aagaccgttctgtgtgtgtgt-tgggtgtgtgtgtgtgtgtgtgtgtgtgt 115
 QY 120 ttgttcttactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 179
 Db 116 -ttcttctgt 173
 QY 180 acaccggagacgactgtgagtgta-aaggatcccttctgtgagacgtgtgt 236
 Db 174 acaccggagacgactgtgagtgtaaaagatccttcttgaagtgatgtgtgt 233
 QY 237 ggcggcgctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 296

Db 234 ggcggcgctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 293
 QY 297 tctcttagacgctgaaagactgtgagactgtgtgtgtgtgtgtgtgtgtgtgt 356
 Db 294 tctcttagacgctgaaagactgtgagactgtgtgtgtgtgtgtgtgtgtgtgt 353
 QY 357 gctgcacccctgt 416
 Db 354 gctgcacccctgt 413
 QY 417 tactgt 476
 Db 414 tactgt 473
 QY 477 tccgactcttctgt 536
 Db 474 tccgactcttctgt 533
 QY 537 gttctgtctgt 596
 Db 534 gttctgtctgt 593
 QY 597 acagtgactacccctttagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 656
 Db 594 acagtgactacccctttagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 653
 QY 657 cataatgt 716
 Db 654 cataatgt 713
 QY 717 ccaacatgt 776
 Db 714 ccaacatgt 773
 QY 777 gtttaggcaatcttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 836
 Db 774 gtttaggcaatcttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 833
 QY 837 cttaagtgagaaattgtgagaaattgtgagaaattgtgagaaattgtgagaaatt 896
 Db 834 cttaagtgagaaattgtgagaaattgtgagaaattgtgagaaattgtgagaaatt 893
 QY 897 ccaagaaagcaggt 956
 Db 894 ccaagaaagcaggt 953
 QY 957 aaagtgtgagcccttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1016
 Db 954 aaagtgtgagcccttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1012
 QY 1017 accccaactgt 1075
 Db 1013 accccaactgt 1072
 QY 1076 cctctctgt 1135
 Db 1073 cctctctgt 1132
 QY 1136 caccctgt 1195
 Db 1133 caccctgt 1192
 QY 1196 gccaggaggcgaattgtgagcagcagcagcagcagcagcagcagcagcagcag 1255
 Db 1193 gccaggaggcgaattgtgagcagcagcagcagcagcagcagcagcagcagcag 1252
 QY 1256 ttggaanaactaacatcccttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1315
 Db 1253 ttggaanaactaacatcccttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1312
 QY 1316 gtcccttaagtcttctcaacagcactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1375
 Db 1313 gtcccttaagtcttctcaacagcactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1372

QY 1376 cttacaaacgagagcgagagagaattctgttagagctagaanaaaatgttccctgggc 1435
Db 1373 cttacaaacgagagcgagagagaattcttagaagctagaanaaaatgttccctgggc 1432
QY 1436 cgaagcgagaccccgagctgtgcaaaatgagatgacatgagattcccttgatccgc 1495
Db 1433 cgaagcgagaccccgagctgtgcaaaatgagatgacatgagattcccttgatccgc 1492
QY 1496 cggttggaactacaacagcgtgagtgagtgagagcttgaaaatctatcgccagctct 1555
Db 1493 cggttggaactacaacagcgtgagtgagtgagagcttgaaaatctatcgccagctct 1552
QY 1556 ggttgcgagcttccggcgccctcaagaagcgcccaattgttgctaaagtaagaggt 1615
Db 1553 ggttgcgagcttccggcgccctcaagaagcgcccaattgttgctaaagtaagaggt 1612
QY 1616 gatcagaagcgaacgaacctccctcgtaattcttgagaggtctcatggaagccttcag 1675
Db 1613 gatcagaagcgaacgaacctccctcgtaattcttgagaggtctcttggaagccttcag 1672
QY 1676 ggggttcaaccttttgatctactacagagagcccaagaagccttcagcttgccctgc 1735
Db 1673 ggggttcaaccttttgatctactacagagagcccaagaagccttcagcttgccctgc 1732
QY 1736 catggcgagctcgagctctggaatcaagaagaaactcaagagacttggaaggttacaaga 1795
Db 1733 catggcgagctcgagctctggaatcaagaagaaactcaagagacttggaaggttacaaga 1792
QY 1796 ggcgtgagttacgtatcgtatgagagagcgagagaagtgatatacgaagaagagacaga 1855
Db 1793 ggcgtgagttacgtatcgtatgagagagcgagagaagtgatatacgaagaagagacaga 1852
QY 1856 agagagagagagaaacgaagaagaagaagagagagaagaagggagaaacgtaag 1915
Db 1853 agagagagagagaaacgaagaagaagaagagagagaagaagggagaaacgtaag 1912
QY 1916 acggaagaagaagaatttgactaaagatcttgccgagctgtgttaagaaggagagacag 1975
Db 1913 acggaagaagaagaatttgactaaagatcttgccgagctgtgttaagaaggagacag 1972
QY 1976 ggaagagagagagagattttaggaataataggtcagagccctagacagtcagaggaactggg 2035
Db 1973 ggaagagagagagagattttaggaataataggtcagagccctagacagtcagaggaactggg 2032
QY 2036 caataagagacccacacgacgaaggaagctgtgctgtatgttaagaagaagaacactggg 2095
Db 2033 caataagagacccacacgacgaaggaagctgtgctgtatgttaagaagaagaacactggg 2092
QY 2096 aaggaactgcccacgaagaagggaaacaaagacggaaggtccttaagcttaagaagaataa 2155
Db 2093 aaggaactgcccacgaagaagggaaacaaagacggaaggtccttaagcttaagaagaataa 2152
QY 2156 agattagagagagagaggttctggaacccctcccgagagcccgaggttaactttgaaagggtga 2215
Db 2153 agattagagagagagaggttctggaacccctcccgagagcccgaggttaactttgaaagggtga 2212
QY 2216 ggggcaaacagttgagttccgtgtgtatgataccgagcgagagatcagtgctgctacaac 2275
Db 2213 ggggcaaacagttgagttccgtgtgtgtatgataccgagcgagagatcagtgctgctacaac 2272
QY 2276 attaggaataactaaagaanaaaalccctgggtgagtgggtgcacagggcaacgagctga 2335
Db 2273 attaggaataactaaagaanaaaalccctgggtgagtgggtgcacagggcaacagctga 2332
QY 2336 tccatgagatcccgagaanaaacggttgaggtggagtgagagcggttaacccctggtttct 2395
Db 2333 tccatgagatcccgagaanaaacggttgaggtggagtgagagcggttaacccctggtttct 2392
QY 2396 ggtacccctgagtgccagtaacccctctaggtagagacttactgacgaagatggagagc 2455
Db 2393 ggtacccctgagtgccagtaacccctctaggtagagacttactgacgaagatggagagc 2452

QY 2456 tcaattctcttgaacaagaagagacgaagtgctgtgtaatacaaaacccatcagct 2515
Db 2453 tcaattctcttgaacaagaagagacgaagtgctgtgtaatacaaaacccatcagct 2512
QY 2516 gttgacccctccaatatgatagtatgataatcagactataatctcccaagtaaacctgata 2575
Db 2513 gttgacccctccaatatgatagtatgataatcagactataatctcccaagtaaacctgata 2572
QY 2576 agatagagctctgtgttgagagagttccccaagccttgagagaacccgagagtggt 2635
Db 2573 agatagagctctgtgttgagagagttccccaagccttgagagaacccgagagtggt 2632
QY 2636 ttgagcaagaagcttccccaaggtta ttaactgaaagcgagtgctacacagatc 2695
Db 2633 ttgagcaagaagcttccccaaggtta ttaactgaaagcgagtgctacacagatc 2692
QY 2696 agtcaagaagctaccccttgatgagaaggtcgaagaagatcttgccgagatccaag 2755
Db 2693 agtcaagaagctaccccttgatgagaaggtcgaagaagatcttgccgagatccaag 2752
QY 2756 attaatccaaagagctatcctgtctgtccaatcccttggaatcccttgatcc 2815
Db 2753 attaatccaaagagctatcctgtctgtccaatcccttggaatcccttgatcc 2812
QY 2816 ggttaagaagccttgagacaaatgatatcgaacagctacagagcttgagagagtcataa 2875
Db 2813 ggttaagaagccttgagacaaatgatatcgaacagctacagagcttgagagagtcataa 2872
QY 2876 aaggtgacagagataatcaccccaagctccgaacccctttaaccccttgagccctcc 2935
Db 2873 aaggtgacagagataatcaccccaagctccgaacccctttaaccccttgagccctcc 2932
QY 2936 ggccttaagaagctgtgtacacagatcttgagacttaaaagaatgctcttcttgctagat 2995
Db 2933 ggccttaagaagctgtgtacacagatcttgagacttaaaagaatgctcttcttgctagat 2992
QY 2996 acaccccaactagccaacacattttgctctgaaatgagagatccaaggtacggagaagc 3055
Db 2993 acaccccaactagccaacacattttgctctgaaatgagagatccaaggtacggagaagc 3052
QY 3056 cgggagagcttcaacttgagacccgaatgctcccaaggttcaagaagctcccgacatcttga 3115
Db 3053 cgggagagcttcaacttgagacccgaatgctcccaaggttcaagaagctcccgacatcttga 3112
QY 3116 cgaagccctcaacagagagacttgcccaacttcaagatccaacacccctcagtgactct 3175
Db 3113 cgaagccctcaacagagagacttgcccaacttcaagatccaacacccctcagtgactct 3172
QY 3176 cgaatagctgagatgacctctcttgcgagagcccaacaaagagctgcttaagaagttac 3235
Db 3173 cgaatagctgagatgacctctcttgcgagagcccaacaaagagctgcttaagaagttac 3232
QY 3236 gaaggaactatctgtggaattgtgctgacctaaggtcacaagcctctgctcaagaagccca 3295
Db 3233 gaaggaactatctgtggaattgtgctgacctaaggtcacaagcctctgctcaagaagccca 3292
QY 3296 gatttgaagagagagtaacatacttggtgtaacagtttgcgggcgagcagatgagct 3355
Db 3293 gatttgaagagagagtaacatacttggtgtaacagtttgcgggcgagcagatgagct 3352
QY 3356 gaagggagcgaggaagaagaacttgaatcagaatagcgcccaacaaacgaagcaagaagt 3415
Db 3353 gaagggagcgaggaagaagaacttgaatcagaatagcgcccaacaaacgaagcaagaagt 3412
QY 3416 gaagagatttttgaggacaagctgattttgacagagttgataccggggttgagcactt 3475
Db 3413 gaagagatttttgaggacaagctgattttgacagagttgataccggggttgagcactt 3472
QY 3476 agcagcccaactacccgctaaacaaagaagaaggagatctccctgggtccctgaagc 3535
Db 3473 agcagcccaactacccgctaaacaaagaagaaggagatctccctgggtccctgaagc 3532
QY 3536 ccagagagcatttgatgatacaaaaagccctgctgagcgacccctgctggccctcc 3595

Db 3533 cccgaagagcatctgacatccaanaaagccctgacgacgacccctgccc 3592
Qy 3596 tgaagtaactaaaccccttaacctatgtgagtagagcgttaagggagtagcccgagagt 3655
Db 3593 tgaagtaactaaaccccttaacctatgtgagtagagcgttaagggagtagcccgagagt 3652
Qy 3656 tttaaccacaaccccttagagacatlgagagagacctgtgtacctacccctgcaaaagacttga 3715
Db 3653 tttaaccacaaccccttagagacatlgagagagacctgtgtacctacccctgcaaaagacttga 3712
Qy 3716 tccctgtagccagtggtgtggtcccatatgacctgaagactatcgagctgtgagcatatggt 3775
Db 3713 tccctgtagccagtggtgtggtcccatatgacctgaagactatcgagctgtgagcatatggt 3772
Qy 3776 caagagccgtgacaattgacctgttgagacagaataaactgtaattagcccccatgacct 3835
Db 3773 caagagccgtgacaattgacctgttgagacagaataaactgtaattagcccccatgacct 3832
Qy 3836 ggaagaacatcgttgcgcagccccacagaccgtatgga tgaaccaagcccgatgacccacta 3895
Db 3833 ggaagaacatcgttgcgcagccccacagaccgtatgga tgaaccaagcccgatgacccacta 3892
Qy 3896 tcaaaagccgtgttccacagagagtgtaacttgcgtccacacgcccgtctcaacctgc 3955
Db 3893 tcaaaagccgtgttccacagagagtgtaacttgcgtccacacgcccgtctcaacctgc 3952
Qy 3956 caacttctgcctgaagagactgatacgaacagtgacatgatgtgacatcaactatgtat 4015
Db 3953 caacttctgcctgaagagactgatacgaacagtgacatgatgtgacatcaactatgtat 4012
Qy 4016 tgaagagactgggtgcgcagaaagacctatacagacataccgctgacgtgagagtgtaac 4075
Db 4013 tgaagagactgggtgcgcagaaagacctatacagacataccgctgacgtgagagtgtaac 4072
Qy 4076 cgtggttaactgacgggaacagctatgtgtgtgaaagttaagagatgagctgtggcgagct 4135
Db 4073 cgtggttaactgacgggaacagctatgtgtgtgaaagttaagagatgagctgtggcgagct 4132
Qy 4136 ggtgtaacgggaccccgacagatctggtgcagacagccctgcggaaagaaattcaagcgaaa 4195
Db 4133 ggtgtaacgggaccccgacagatctggtgcagacagccctgcggaaagaaattcaagcgaaa 4192
Qy 4196 ggtgtagctcatggtccctacgcgaagcttgcgtggtgcgcgaaggaatccataaacat 4255
Db 4193 ggtgtagctcatggtccctacgcgaagcttgcgtggtgcgcgaaggaatccataaacat 4252
Qy 4256 ttatacgacagcaggtatgacctgtgcacacgtatcacgggggccaatcataaaca 4315
Db 4253 ttatacgacagcaggtatgacctgtgcacacgtatcacgggggccaatcataaaca 4312
Qy 4316 aaggggtgtgcttaacctagcagggaggaagaaataaagaacaaaggaatcttaaacct 4375
Db 4313 aaggggtgtgcttaacctagcagggaggaagaaataaagaacaaaggaatcttaaacct 4372
Qy 4376 attgaagacctatattgtgccaanaaggttagctatatacatgctgtcgtgacatacga 4435
Db 4373 attgaagacctatattgtgccaanaaggttagctatatacatgctgtcgtgacatacga 4432
Qy 4436 agccaagaatctatattagaggaacacagatggtcgtgacccgggttgcgaagcagcagc 4495
Db 4433 agccaagaatctatattagaggaacacagatggtcgtgacccgggttgcgaagcagcagc 4492
Qy 4496 ccaaggtgttlaaccttgcctcaataaagaacgccaagcccccaagcccccaagcagaca 4555
Db 4493 ccaaggtgttlaaccttgcctcaataaagaacgccaagcccccaagcccccaagcagaca 4552
Qy 4556 gtaacaccttagaagactggtgcaagaagataaagaatagacacgttctctgagactccga 4615
Db 4553 gtaacaccttagaagactggtgcaagaagataaagaatagacacgttctctgagactccga 4612
Qy 4616 gggagacgtgctatctcatatgtggaaggaatctgtccccaacaaagaggttagata 4675

Db 4613 agggacgtgtatacctcaga tgggaaggaatccctgtccccaacaaagaggttagata 4672
Qy 4676 tgtccaacagatatacgttcttaacccacttagagaaactaaacaccttgacgaagtgtgcag 4735
Db 4673 tgtccaacagatatacgttcttaacccacttagagaaactaaacaccttgacgaagtgtgcag 4732
Qy 4736 aacatcccttatatgttctgaagcttaacggaggtgtgcagactcgtgtgtcaacaatg 4795
Db 4733 aacatcccttatatgttctgaagcttaacggaggtgtgcagactcgtgtgtcaacaatg 4792
Qy 4796 tgtcccttgccagctgtgttaagtactaactccctccagaatacctccagaagaagactaaag 4855
Db 4793 tgtcccttgccagctgtgttaagtactaactccctccagaatgctccagaaggagagactaaag 4852
Qy 4856 gggagaaccccaagcgctacgttgaggaaggtgacttcaacgtgaggttaagcccgctaaata 4915
Db 4853 gggagaaccccaagcgctacgttgaggaaggtgacttcaacgtgaggttaagcccgctaaata 4912
Qy 4916 cggaaacaaatatactatgttcttgaagacaccttccaggtatgggttagagcttacc 4975
Db 4913 cggaaacaaatatactatgttcttgaagacaccttccaggtatgggttagagcttacc 4972
Qy 4976 tactaagaagaagacttcaacgtgtgtgcctaagaanaatactggaggaataatttccaaag 5035
Db 4973 tactaagaagaagacttcaacgtgtgtgcctaagaanaatactggaggaataatttccaaag 5032
Qy 5036 attggaaatcccttaaggttaaggttgaagacaatggtccagcttgcgttgcacaggttaag 5095
Db 5033 attggaaatcccttaaggttaaggttgaagacaatggtccagcttgcgttgcacaggttaag 5092
Qy 5096 tcaagggactgtgccaaatcttggggatgtgttggaaactgcatgtgtcatatagaagaccca 5155
Db 5093 tcaagggactgtgccaaatcttggggatgtgttggaaactgcatgtgtcatatagaagaccca 5152
Qy 5156 aagctcagagcaggttgaagagagatgataaaccatataaagaaccccttaactaatgac 5215
Db 5153 aagctcagagcaggttgaagagagatgataaaccatataaagaaccccttaactaatgac 5212
Qy 5216 cacagaagactggtataatgttgatggtctccctgcgacctgtgtctttttaggttgaag 5275
Db 5213 cgcggagacgtggtgttaatgtatctgtagactccctgcgacctgtgtctttttaggttgaag 5272
Qy 5276 gaacacccctggaacagtttgcgtgacccccctatgatatgtcttaacggggagacccccc 5335
Db 5273 gaacacccctggaacagtttgcgtgacccccctatgatatgtcttaacggggagacccccc 5332
Qy 5336 gttggcagaataatgacctgtgcacatagtgctgtgtgtgtgtcttcccaagccttgtctc 5395
Db 5333 gttggcagaataatgacctgtgtgcacatagtgctgtgtgtgtgtgtcttcccaagccttgtctc 5392
Qy 5396 taaggtcaagcgctcgaagtggtgtgagcgacggtggaagcagctccgggaagccta 5455
Db 5393 taaggtcaagcgctcgaagtggtgtgagcgacggtggaagcagctccgggaagccta 5452
Qy 5456 ctc---aggaagagacttgcgaatctccacatcgtctcccaagttgagatctacgtatgt 5512
Db 5453 ctcgaagagggagacttgcgaatctccacatcgtctcccaagttgagatctacgtatgt 5512
Qy 5513 taagaagcagcgtgcaagaaacctcgaagacgtgtggaagggagacattatcgtacttt 5572
Db 5513 taagaagcagcgtgcaagaaacctcgaagacgtgtggaagggagacattatcgtacttt 5572
Qy 5573 gaccacacaaagcggtgttaaaatgtgaagaagaatccccacttgaatccatgcatcccaagt 5632
Db 5573 gaccacacaaagcggtgttaaaatgtgaagaagaatccccacttgaatccatgcatcccaagt 5632
Qy 5633 taagccggcgccacccctccagactcgggttgaagagcggaaagagactggaatccctttaa 5692
Db 5633 taagccggcgccacccctccagactcgggttgaagagcggaaagagactggaatccctttaa 5692
Qy 5693 gcttgcctccatcgtcgtgttccctactctcaataaatactcccaagcgagtagtaaac 5752
Db 5693 gcttgcctccatcgtcgtgttccctactctcaataaatactcccaagcgagtagtaaac 5752

QY	5753	gcctatagacagctgaaaccccatagaccttattcccttaoctggtatattgacc	5812
Db	5753	gcataagagagacccctgaatcccatatacccttattctcactggttaattactgact	5812
QY	5813	ctgataagaggtgtacacgtataatagcctgaggtgtgtctcttaagagcactgtgtgc	5872
Db	5813	ccggcacaggtattataatacaaacactcaaggggaggtcctctttagagacctgtgtgc	5872
QY	5873	ctgaactgcatlctcgtccctcogattgatacccgctgtttaaagcacactcccaac	5932
Db	5873	ctgactcatacgtttgtcctcagatcagttatc---ctagctcgaactcccccagata	5929
QY	5933	tagcctgattgattggttctatgtgtcccaag---sacagagaagaagaatactgtg	5989
Db	5930	tcctcactgtcagcgatttattatgtgtcccaagacacacaataatggaataactgtgc	5989
QY	5990	ggaggtcttgaggaaatcctctgttagagatgaggtgtgtcactcactccacgaltgagact	6049
Db	5990	gaatcccaagagattcttctgttaacaaatggaactgtgtaacctcaactgactgattt	6049
QY	6050	ggaaatggccgactctctcctcagacccgggttaaatctcctctgttaa-----	6098
Db	6050	ggaaatggccgactctcctcagacccgggttaaatctcctctgttaa-----	6098
QY	6099	-ttccggcccgagcaaglacaaatgataacataataagataagagctgtcccat	6157
Db	6110	gctctgagcaaatattactactgacccgtgattgaaactggaagcccaagtgtcctctc	6169
QY	6158	cagaactagattatcctaagaataagtttact--gaaaggaanaacaggaataattc	6215
Db	6170	cagacttagattatcctaagaataagtttact--gaaaggaanaacaggaataattc	6229
QY	6216	agtgatgaataatggtatagcgtgaggaataatttttattatgtgcggggagag--cag	6271
Db	6230	aatggatgaataatggtatagcgtgaggaataatttttattatgtgcggggagag--cag	6289
QY	6272	ggtcacttttaacattcgtccttagagacagggagacagaaacccctgttgcaattg	6331
Db	6290	gtccatttcaactatctgcctcaataataaacagctg---gagcctccaaatgtctatg	6346
QY	6332	gaccggataaagctactgtgtgaacaggggccccggccctgagacacacgataactgtc	6391
Db	6347	gaccataaactgcttgaacgggtcaagaagcccccaacaggaacgagacatcctct-	6405
QY	6392	cggtgcccccaataacactcgtcgtcgccgtgacataacacagccgctagacaacgatac	6451
Db	6406	-----aactaactc	6415
QY	6452	ctgattgattcctacacaacgactagaaactcccaaggtgtcctgttaagaacagac	6511
Db	6416	ctgattgattcctacacaacgactagaaactcccaaggtgtcctgttaagaacagac	6460
QY	6512	agagactccttagctctacacagggaggttcccaagacatcaactcagacacgactgtg	6571
Db	6461	caaaactttagcctcactcagagaggttcttaagcttcaactcagacacgactcctcag	6520
QY	6572	ccacttctctgtgtgtgtgtctatcctcagggccctccttattatgagggaggtgcta	6621
Db	6521	taactcctctgt	6580
QY	6632	aagaaagaanaattcaatgtgacaaagacatagaatcaatgactatgaggtgtccgaa	6691
Db	6581	gaagggggaattcaatgtgacaaagacatagaatcaatgactatgaggtgtccgaa	6640
QY	6692	ataagcttaacctcaatgaatttcccggaagggagacatgatatggaanaactcccat	6751
Db	6641	ataagcttaacctcaatgaatttcccggaagggagacatgatatggaanaactcccat	6700
QY	6752	ccacacacacactgtgactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	6811
Db	6701	ccacacacacactgt	6760

QY	6812	tagtactcgtttatacagaaggtgtgtgcatgacatgactggtgttaacccccgtgttcca	6871
Db	6761	tggtactcgtttatagacaggtgtgtgtgcatgacatgactggtgttaacccccgtgttcca	6820
QY	6872	cctcagcttcaacaaatccaaagaattctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	6931
Db	6821	cctcagcttcaacaaatccaaagaattctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	6880
QY	6932	actcactccttaggaaggt	6991
Db	6881	actcactccttaggaaggt	6940
QY	6992	gagaaacgctaccccttaacactagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	7051
Db	6941	gagaaacgctaccccttaacactagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	7000
QY	7052	gaacagggacacactgcccctgatacagaagccacacacacacacacacacacacacac	7111
Db	7001	gaacagggacacactgcccctgatacagaagccacacacacacacacacacacacacac	7060
QY	7112	tacatgcgctcagacagaagatctccgagccttaagagaggtgtgtgtgtgtgtgtgtgt	7171
Db	7061	tacatgcgctcagacagaagatctccgagccttaagagaggtgtgtgtgtgtgtgtgtgt	7120
QY	7172	agtcctcagactcctgt	7231
Db	7121	agtcctcagactcctgt	7180
QY	7232	tcttaagaagaaggt	7291
Db	7181	tcttaagaagaaggt	7240
QY	7292	actcagagac	7351
Db	7241	actcagagac	7300
QY	7352	gggaaagaagaaggt	7411
Db	7301	gggaaagaagaaggt	7360
QY	7412	ccacactccttctgt	7471
Db	7361	ccacactccttctgt	7420
QY	7472	gacctgtgttaataataggt	7531
Db	7421	gacctgtgttaataataggt	7480
QY	7532	tcatgttacttaggacaggt	7591
Db	7481	tcatgttacttaggacaggt	7537
QY	7592	cttccagctcctaagattgaactataacagaacagaagaggtgtgtgtgtgtgtgtgtgt	7651
Db	7538	cttccagctcctaagattgaactataacagaacagaagaggtgtgtgtgtgtgtgtgtgt	7597
QY	7652	aatgcacaactaacctccagaac	7676
Db	7598	aatgcacaactaacctccagaac	7622

RESULT	5
AAC67023	
ID	AAC67023 standard; DNA; 7873 BP.
XX	
AC	AAC67023;
XX	
DT	27-MAR-2001 (first entry)
XX	
DE	PERV env protein coding sequence SEQ ID NO: 23.
XX	
KW	Xenotransplantation; infectious agent; vaccine; ds.
XX	

Db 1620 aagaagataaagactaaggaagacggggttcggagacccctcccgagccaggttaactt 1679
QY 2206 tgaaggtgagaggggcaaccaggttgagttcctgtgtatccggagcggaacttcagtc 2265
Db 1680 tgaaggtgagaggggcaaccaggttgagttcctgtgtatccggagcggaacttcagtc 1739
QY 2266 tgcatacaacactaagaaactaaagaaataatcctgtgtatccggagcggaacttcagtc 2325
Db 1740 tactacagccactaagaaactaaagaaataatcctgtgtatccggagcggaacttcagtc 1799
QY 2326 aacggcgatccatgagactacccgaagacggttgacttggaggtgagcggttaacc 2385
Db 1800 aacacagatccatgagactacccgaagacggttgacttggaggtgagcggttaacc 1859
QY 2386 actggtttgtatccctgaggtgaggttacccttcctgaggtgaggttaccctgag 2445
Db 1860 actggtttgtatccctgaggtgaggttacccttcctgaggtgaggttaccctgag 1919
QY 2446 agatggagagctcaaatctcttgaacaaggaagacagagtgctgtgataaacaac 2505
Db 1920 agatggagagctcaaatctcttgaacaaggaagacagagtgctgtgataaacaac 1979
QY 2506 ccaatcactgtgtgagccctcccaatagatgataatccctacatattccccaagtaa 2565
Db 1980 ctatccactgtgtgagccctcccaatagatgataatccctacatattccccaagtaa 2039
QY 2566 agctgtacagatatacagctcgtgtgagacggttccccaagcctggcggaagacg 2625
Db 2040 agctgtacagatatacagctcgtgtgagacggttccccaagcctggcggaagacg 2099
QY 2626 caaggtatgggttggcaaaagcaagttccccaagcgttatcaactgaagcgagtgcta 2685
Db 2100 caaggtatgggttggcaaaagcaagttccccaagcgttatcaactgaagcgagtgcta 2159
QY 2686 caacagatcagtcagacagatcccttgagtagagaggtcggaaggaattggcgcc 2745
Db 2160 caacagatcagtcagacagatcccttgagtagagaggtcggaaggaattggcgcc 2219
QY 2746 atgttcaagaattatcccaagagcagctagcttcctcccaactccctggagatc 2805
Db 2220 atgttcaagaattatcccaagagcagctagcttcctcccaactccctggagatc 2279
QY 2806 cccgtgtaacggttagaagagcctgggacaaatgatacgaacagtaacagacttgaag 2865
Db 2280 cccgtgtaacggttagaagagcctgggacaaatgatacgaacagtaacagacttgaag 2339
QY 2866 aggtcaataaagaggtgagacatacccaacggtcccgaaacctataactcttga 2925
Db 2340 aggtcaataaagaggtgagacatacccaacggtcccgaaacctataactcttga 2399
QY 2926 ggcgcctcccgctgaagagcgtgatacagatctggaacttaagaatgacctctc 2985
Db 2400 ggcgcctcccgctgaagagcgtgatacagatctggaacttaagaatgacctctc 2459
QY 2986 gctctgagattacaccccaatagccaacacttttgccttcgaaatgagagatccagta 3045
Db 2460 gctctgagattacaccccaatagccaacacttttgccttcgaaatgagagatccagta 2519
QY 3046 cgggaagaacccgggagactacccctggagccgagctggcccaaggttcaagaactcccg 3105
Db 2520 cgggaagaacccgggagactacccctggagccgagctggcccaaggttcaagaactcccg 2579
QY 3106 ccatctttgagagagccttacacaggaacttggccaactltagatccaaacccctag 3165
Db 2580 ccatctttgagagagccttacacaggaacttggccaactltagatccaaacccctag 2639
QY 3166 tgaacctccctacagtgagtgagacgttcttggcgagagccaacaaacagagactgt 3225
Db 2640 tgaacctccctacagtgagtgagacgttcttggcgagagccaacaaacagagactgt 2699
QY 3226 tagaagtagaagagcactactgtgagattgtctgagactaggtacagagccttgcta 3285
Db 2700 tagaagtagaagagcactactgtgagattgtctgagactaggtacagagccttgcta 2759
QY 3286 agaagccagatttgcagagagaggttaacatacttgggtgagacttgcggcgag 3345
Db 2760 agaagccagatttgcagagagaggttaacatacttgggtgagacttgcggcgag 2819
QY 3346 agcgtatgctgagcagagcagcgaagaaactgtatccagaatcccgcccaacag 3405
Db 2820 agcgtatgctgagcagagcagcgaagaaactgtatccagaatcccgcccaacag 2879
QY 3406 ccaaaaagtggagaggttttgggagacggtgttggagacttggagactccggggt 3465
Db 2880 ccaaaaagtggagaggttttgggagacggtgttggagacttggagactccggggt 2939
QY 3466 ttgcaccttaagcagcccaactcccgctaaccaagaagaaagggggttccctgg 3525
Db 2940 ttgcaccttaagcagcccaactcccgctaaccaagaagaaagggggttccctgg 2999
QY 3526 ctccctgagacagaagagcatttgatatacaaaagccctgtgagcgacctgtc 3585
Db 3000 ctccctgagacagaagagcatttgatatacaaaagccctgtgagcgacctgtc 3059
QY 3586 tggccctccctgagcagtaacaccccttaaccccttatgtgagtagagcgaagtag 3645
Db 3060 tggccctccctgagcagtaacaccccttaaccccttatgtgagtagagcgaagtag 3119
QY 3646 cccgagagattttaaaccacaccccttaggacaaatggagagacgttgcctacgtgaa 3705
Db 3120 cccgagagattttaaaccacaccccttaggacaaatggagagacgttgcctacgtgaa 3179
QY 3706 agaagcttatccctgtagccaggtgtgtgagcccgatgtgagagctatccgagctgtg 3765
Db 3180 agaagcttatccctgtagccaggtgtgtgagcccgatgtgagagctatccgagctgtg 3239
QY 3766 ccaactgtgtcgaagagcgttgcgaatgacttggagagagataaactgtatagccc 3825
Db 3240 ccaactgtgtcgaagagcgttgcgaatgacttggagagagataaactgtatagccc 3299
QY 3826 cccatgcatltggagaaacatcgttgcgacgcccacgaacgagatgagatgacacgcccga 3885
Db 3300 cccatgcatltggagaaacatcgttgcgacgcccacgaacgagatgagatgacacgcccga 3359
QY 3886 tgaaccataataaagcgtcttccacagaagaggtcaacttgcctccacagccgtc 3945
Db 3360 tgaaccataataaagcgtcttccacagaagaggtcaacttgcctccacagccgtc 3419
QY 3946 tcaaccttcgaacttctgcctggaagagacgtatgaacagatgagatgacatc 4005
Db 3420 tcaaccttcgaacttctgcctggaagagacgtatgaacagatgagatgacatc 3479
QY 4006 aactatgtatgagagacgttgggttcgcaagacctaacaacataccgctgagtag 4065
Db 3480 aactatgtatgagagacgttgggttcgcaagacctaacaacataccgctgagtag 3539
QY 4066 aagtgctaacctgtgtcactgagcaggaagcgtatgtgtggaaggtgaagagtgctg 4125
Db 3540 aagtgctaacctgtgtcactgagcaggaagcgtatgtgtggaaggtgaagagtgctg 3599
QY 4126 gggcgagcagtggtgagcaggaacccgagcagatctgggcccagcgttcggaaggaact 4185
Db 3600 gggcgagcagtggtgagcaggaacccgagcagatctgggcccagcgttcggaaggaact 3659
QY 4186 cagcgcaaaagcctgagcctacatggccctcaagcagcttggcgctggcggaaggaat 4245
Db 3660 cagcgcaaaagcctgagcctacatggccctcaagcagcttggcgctggcggaaggaat 3719
QY 4246 ccaataacattatcagcagcaggtatgtccttggcgactgacacgtatccagggcca 4305
Db 3720 ccaataacattatcagcagcaggtatgtccttggcgactgacacgtatccagggcca 3779
QY 4306 tctataaacaaggggttggcttaacctgagcagggggaagaaataagaacaaggggaa 4365
Db 3780 tctataaacaaggggttggcttaacctgagcagggggaagaaataagaacaaggggaa 3839

Db 5877 cagtcaccactga-----gcctaacatcactattataaaacagggcgaaac 5921
QY 6518 tottaagttctccaggaggtttcccaagccatcaactccaccgacctgatgcactt 6577
Db 5922 tttttaacctccaggaggtttcccaagctcttaactccaccgacctccagagctacct 5981
QY 6578 ctctctgttggtttgtctatctccagggcctcttattatagagggttgcttaaaagaa 6637
Db 5982 ctctctgttggtttgtctatctccagggcctcttattatagagggttgcttaaaagaa 6641
QY 6638 gaaatctcaatgtgaccagaagcagataaatacaatgtacatgggggtcccgaaataagc 6697
Db 6042 ggaatctcaatgtgaccagaagcagataaatacaatgtacatgggggtcccgaaataagc 6101
QY 6698 ttaccctactgaagtgttcctgggaaggagacatgcataaggaaagctcccccaccacc 6757
Db 6102 ttaccctactgaagtgttcctgggaaggcactgcataggatgttccccaccaccacc 6161
QY 6758 aacacctttgtatagtagtctgtttatagcaggccctcgaaataatagtagttatag 6817
Db 6162 aacacctttgtatagtagtctgtttatagcaggccctcgaaataatagtagttatag 6221
QY 6818 ctggtttaaacaggtgtggcagtcacatactgggttaaacccctgtgtttccacctcag 6877
Db 6222 ctggtttaaacaggtgtggcagtcacatactgggttaaacccctgtgtttccacctcag 6281
QY 6878 tottaaccacatccaaagatttctgtctatgttccaaatcgtcccccagagtgactacc 6937
Db 6282 ttttaaccacatccaaagatttctgtctatgttccaaatcgtcccccagagtgactacc 6341
QY 6938 atcttgaggaggtgtctctgataatgataatgataatgataatgataatgataatgata 6997
Db 6342 atcttgaggaggtgtctctgataatgataatgataatgataatgataatgataatgata 6401
QY 6998 cgtatcccttaccctagctatgataatgataatgataatgataatgataatgataatgata 7057
Db 6402 cgtatcccttaccctagctatgataatgataatgataatgataatgataatgataatgata 6461
QY 7058 ggaagctgcctctgatacagagacacagcagctagagaaaggactgtgtgagctacatg 7117
Db 6462 gaaagctgcctctgatacagagacacagcagctagagaaaggactgtgtgagctacatg 6521
QY 7118 cggccatgacagaagatctccagccttaaaaggaggtgtgtgacaaactgaagagctccc 7177
Db 6522 gaattgaacggaagatctccagccttaaaaggaggtgtgtgacaaactgaagagctccc 6581
QY 7178 tgaactctcttctgaagtggtctacagaaacccgaggaatgtatagatctgtttcttaa 7237
Db 6582 taactctcttctgaagtggtctacagaaacccgaggaatgtatagatctgtttcttaa 6641
QY 7238 gagaaggtgggttatgtgagccttaaaaggaggtgtgtctctatgtatagatcactcag 7297
Db 6642 aagaagaggattatgtgagccttgaaggaggaatgtgtttttatgtgagatcactcag 6701
QY 7298 gagccatcagagactcactgaacaaagcttgaaaaaa----- 7334
Db 6702 gggccatcagagactcactgaacaaagcttgaaaaaa----- 7334
QY 7335 ----- 7334
Db 6762 cgaacacccagacaggtgtctgttttcaataaagaactgtcgaaagtaaacctgagct 6821
QY 7335 -----gttagagagcgtcgaaagggaagaaagagaggtgacccaggggt 7375
Db 6822 aaagccagcttagagaaggtgtgagagcgtcgaaagggaagaaactactcaagagt 6881
QY 7376 ggtttgaaggatggttcaaacaggtctccttgatgacacacccctgtcttctgtctacgg 7435
Db 6882 ggtttgaaggatggttcaaacaggtctccttgatgacacacccctgtcttctgtctacgg 6941
QY 7436 gggcccttagtct 7495
Db 6942 gacccttaaatgct 7001

QY 7496 ttgctttgttagagaacaggtgagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcag 7555
Db 7002 ttgctttgttagagaacaggtgagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcag 7061
QY 7556 aaggccttctgagcccaaggagaacagtcagtcagtcagtcagtcagtcagtcagtcagtcag 7615
Db 7062 aaggccttctgagcccaaggagaacagtcagtcagtcagtcagtcagtcagtcagtcagtcag 7118
QY 7616 attaaacaagaagaagtgagggaatgaaagatgaaatgaaatgaaatgaaatgaaatgaaatgaa 7675
Db 7119 attaaacaagaagaagtgagggaatgaaagatgaaatgaaatgaaatgaaatgaaatgaaatgaa 7178
QY 7676 ccagggaagttaataaaagctctaaatgcgcccgaaattacagacctgctggctgccagt 7735
Db 7179 ccagggaagttaataaaagctctaaatgcgcccgaaattacagacctgctggctgccagt 7232
QY 7736 aatataggagaaggtcacacttctctattgttccaggccctgtctctctctctctctctctctct 7795
Db 7233 --ataggtaaaagatcactctttgtctgttttagggtgtctctctctctctctctctctctct 7290
QY 7796 ataacaggaaatgagttga----- 7814
Db 7291 ctttgggaagggaagaaacaggccctcagtgatgtgcctctctctctctctctctctctctct 7350
QY 7815 -----ctaactcgtcttctgattctgttaaacctgactg----- 7848
Db 7351 aactgctcctaaactgctgtttgtctctgttaaacctgctgtgcataagataaaagagga 7410
QY 7849 -----gcaccataagaagaattgattacacattgacacccctctagtgacc 7891
Db 7411 gaagtcaattgcctaaacgagcccccagtaagatcggtgtaccacaaaatgttgaaacaca 7470
QY 7892 tatctcaactgcaatgtcactctctccaggagccacagcagatgcggacctccggagc 7951
Db 7471 tatctgttgacaaactgtctcccaccccccgaacacatgcgaaatgtgtaactctctctctct 7530
QY 7952 tattttaaaatgattgtccagggagcgcggtctctcgatatttttaaaatgattgtgtcca 8011
Db 7531 caatttaaaatgattgtccagggagcgcggtctctcgatatttttaaaatgattgtgtctct 7587
QY 8012 tggagcgcgggtctctcgatatttttaaaatgattgtgtgtgacgcacagcgtgtgtgtg 8071
Db 7588 -----ttgtgattttgaaatgattgtgtgtgaaagcggtgtgtgtgtgtgtgtgtgtgt 7635
QY 8072 aaccccataaaagctgtcccgactcccgactccggggcgagcactctctacccctcggtg 8131
Db 7636 aaccccataaaagctgtcccgactcccgactccggggcgagcactctctacccctcggtg 7695
QY 8132 tgcagcactgtggccccagcgcttggaataaaaaatcctctctgtgtgtgtgtgtgtgtgtgtgt 8191
Db 7696 tgcagcactgtggccccagcgcttggaataaaaaatcctctctgtgtgtgtgtgtgtgtgtgtgt 7755

RESULT 6

AAC67019
ID AAC67019 standard; DNA; 7362 BP.

XX AAC67019;

XX 27-MAR-2001 (first entry)

XX PERV env protein coding sequence SEQ ID NO: 19.

XX Xenotransplantation; infectious agent; vaccine; ds.

XX Porcine endogenous retrovirus.

XX WO200071726-A1.

XX 30-NOV-2000.

XX 24-MAY-2000; 2000WO-US14296.

QY	3956	cactcttgcctgaagagactgatgaaccagtgactcatgatgtgccatcaactattgat	4015
Db	4616	cactcttgcctgaagagactgatgaaccagtgactcatgatgtgccatcaactattgat	4675
QY	4016	tgaagagactgggtccgcaaggaaccttacagacataccgtgactgagaagaagtctaac	4075
Db	4676	tgaagagactgggtccgcaaggaaccttacagacataccgtgactgagaagaagtctaac	4735
QY	4076	ctgtgtcaactgaaggagcagctatgtgtgaaaggttaagaggatggctgggcgcgcagt	4135
Db	4736	ctgtgtcaactgaaggagcagctatgtgtgaaaggttaagaggatggctgggcgcgcagt	4795
QY	4136	ggttgagcggaccgcacgatctggccagcagctgcgggaaggaacttcagcgcacaaa	4195
Db	4796	ggttgagcggaccgcacgatctggccagcagctgcgggaaggaacttcagcgcacaaa	4855
QY	4196	ggctgagctcatggcctcacgaagctttgcgctggccgaagggaattccataaacat	4255
Db	4856	ggctgagctcatggcctcacgaagctttgcgctggccgaagggaattccataaacat	4915
QY	4256	ttatacggacagcagtgatgcctttgcgactgcacacgtcacacgggccatctataaaca	4315
Db	4916	ttatacggacagcagtgatgcctttgcgactgcacacgtcacacgggccatctataaaca	4975
QY	4316	aagggtgtgtcttacctcagcaggagggaataaagacaaagaggaaattcttaagcct	4375
Db	4976	aagggtgtgtcttacctcagcaggagggaataaagacaaagaggaaattcttaagcct	5035
QY	4376	attagaagccttacatttgcgaagagcgtactattatacacgtgctcctggacatcagaa	4435
Db	5036	attagaagccttacatttgcgaagagcgtactattatacacgtgctcctggacatcagaa	5095
QY	4436	agccaaagatctcatctatagagggaacacagatggctgacccgggttcgcaagcaggcagc	4495
Db	5096	agccaaagatctcatctatagagggaacacagatggctgacccgggttcgcaagcaggcagc	5155
QY	4496	ccagggtgttaacctctgcctataatagaacccccaagcccaagcccaagcccaagcagca	4555
Db	5156	ccagggtgttaacctctgcctataatagaacccccaagcccaagcccaagcccaagcagca	5215
QY	4556	gtacaccttagaagactggcgaagagataaaaagatagacagttctctgagactccgga	4615
Db	5216	gtacaccttagaagactggcgaagagataaaaagatagacagttctctgagactccgga	5275
QY	4616	ggggacctgctatacctcatatgggaaggaatacctgccccacaaagaggttagaata	4675
Db	5276	agggacctgctatacctcatatgggaaggaatacctgccccacaaagaggttagaata	5335
QY	4676	tgtccacagatatactgtctaacccactagggaactaaacacctgacaggttgctcag	4735
Db	5336	tgtccacagatatactgtctaacccactagggaactaaacacctgacaggttgctcag	5395
QY	4736	aacatcccttatactgttctgaggtaccaggagtggtgactcgggtggtcaaacatttg	4795
Db	5396	aacatcccttatactgttctgaggtaccaggagtggtgactcgggtggtcaaacatttg	5455
QY	4796	tgtgccctgccagctggttaatgtctaactctccagataactccagaaagagactaag	4855
Db	5456	tgtgccctgccagctggttaatgtctaactctccagataactccagaaagagactaag	5515
QY	4856	gggaagcccccagcgcctcactgggaagtggacttcactgaggtaaagccggctaaata	4915
Db	5516	gggaagcccccagcgcctcactgggaagtggacttcactgaggtaaagccggctaaata	5575
QY	4916	cggaaacaataatctattgggtttttgtagacaccttttcaggatgggtgaggttattcc	4975
Db	5576	cggaaacaataatctattgggtttttgtagacaccttttcaggatgggtgaggttattcc	5635
QY	4976	tactaagaagagacttcaaccgttgggtgctaagaataactggaggaatttttccaaag	5035
Db	5636	tactaagaagagacttcaaccgttgggtgctaagaataactggaggaatttttccaaag	5695

QY	5036	atttggataactaaaggttaataagggtcagacaataaggtccagcttctcgttgccaggttaag	5095
Db	5696	atttggataactaaaggttaataagggtcagacaataaggtccagcttctcgttgccaggttaag	5755
QY	5096	tcaaggactggccaagatatgtggggatttggaattggaactgcaattgtgcatacagacccca	5155
Db	5756	tcaaggactggccaagatatgtggggatttggaattggaactgcaattgtgcatacagacccca	5815
QY	5156	aagctcaggacagatgagagagatgaatagaaccattaaagagacccctaccacaattgac	5215
Db	5816	aagctcaggacagatgagagagatgaatagaaccattaaagagacccctaccacaattgac	5875
QY	5216	cacagagactggcattaatgatgtggatggctctcgtcctcgtgcctttgtgctttttagggtag	5275
Db	5876	cgggagacttgggttaattgatgtggatggctctcgtcctcgtgcctttgtgctttttagggtag	5935
QY	5276	gaacacccctggacagtttggtctgaacccctatgaattgtctacggggagacccccccc	5335
Db	5936	gaacacccctggacagtttggtctgaacccctatgaattgtctacggggagacccccccc	5995
QY	5336	gttggcgaaaattgcctttgtcacatatgtctgatgtgctgctttcccaagcctttgtctc	5395
Db	5996	atgttgaaaattgcttcogtacatatgtctgacgtgctgctttcccaagcctttgtctc	6055
QY	5396	taggtctaaaggcgtcgcagttgggtggagcgagcgagctgggaagcagctccggggagccta	5455
Db	6056	taggtctaaaggcacttgagttgggtgagacaacgagcgtggaggcaactccggggagccta	6115
QY	5456	ctc---aggaggagacttgcaagttccacatcgcttccaaagtggagagattcagtcctatgt	5512
Db	6116	ctcaggaggaggagacttcagatcccatcgcttccaaagtggagagattcagtcctatgt	6175
QY	5513	tagacgccacgtgtgcaggaaaactcgagactcgtgtggaggagacctaactctcgtactttt	5572
Db	6176	tagacgccacgtgtgcaggaaaactcgagactcgtgtggaggagacctaactctcgtactttt	6235
QY	5573	gaccacacaaacgctgtgaaagtgcgaaggaaatccccaccctggatcgcataccacgt	5632
Db	6236	gaccacacaaacgctgtgaaagtgcgaaggaaatccccaccctggatcgcataccacgt	6295
QY	5633	taagcggcgccaactcccgactcgggtgtgagagaccgaaagaagactgagaaatccccctaa	5692
Db	6296	taagcggcgccaactcccgactcgggtgtgagagaccgaaagaagactgagaaatccccctaa	6355
QY	5693	gcttcgctccactcgcctggttccttactcttaacataaactccccagggcagtagtaaac	5752
Db	6356	gcttcgctccactcgcgtggttccttactctgtcaataactcctcaggttaaaggtaaac	6415
QY	5753	gccttatagacagctcgaaccccccatagacattatcccttaccctggctgattattgacc	5812
Db	6416	gccttgtggacagccgaaactccataaaccttatctcactcgtgttacttactgact	6475
QY	5813	ctgatacgggtgtcactgttaatatgactcaggtgtgtgtcctcctagaggcaccttggtcgc	5872
Db	6476	ccggtacaggtattaaatatcaacgactcgaaggggaggtcccttggggaccttggtcgc	6535
QY	5873	ctgaactgattctgcctccagattgattaacccgcgttta-----aagcacacctc	5926
Db	6536	ctgaattatgtctgcctctcagtaacatcccttgcctcagtgctcaatgacagggcacccccc	6595
QY	5927	ccaacctagtcogtagttatgggtttctattgctgccagg---cacagagaaagagaaat	5983
Db	6596	ccgattactcgtcttacgsgttttacgtttgtcccaagaccccccaataatgaagaat	6655
QY	5984	acttggggggttctggggaatcctctctgtaggagatggagctcgctcacctccaaagatg	6043
Db	6656	atttggaaaactcagggaattcttttgcagagcaatggagctcgtaactctctaaatgcatg	6715
QY	6044	gagactggaaatggcgaatctctcctccagagaccgggttaaaattctcctttgt-----	6095
Db	6716	ggaattggaaatggccagtcctcctcagcaagacagatgaagttaactcttttgttaacaaatc	6775
QY	6096	-----caattccgcgcgggcaag	6114

Db 6776 ctaccagttataatcaattattggtccatggagagattggcaacagcggg 6835
 QY 6115 tacaatagatgaataataaagataagagctgctcccatcagacttagattatcta 6174
 Db 6836 tacaataaagatg-tacgaataaagcaataaagctgtcatctgttagacctagattactta 6894
 QY 6175 aagataagtttcaactg--aaaggaaacagagaaatattcaaaagtggataaaatggatg 6232
 Db 6895 aaaaataagtttcaactgaaagaaagaaacagaaatattcaaaagtggtaaaatggatg 6954
 QY 6233 acgtgggaataagttttttat----tatatggcggggagagaggtccacttttaacatt 6288
 Db 6955 tcttgggaataagtttactatggaggtcctggagaaagaaagatctgttctgactatt 7014
 QY 6289 qccttagatagagacgggagacagaccctgtggaatggagcccgataaaagtactg 6348
 Db 7015 gcctcagaataagaactcagatggaacctcgggtgctatagaccataaaggtttg 7074
 QY 6349 gctgaacagggggcccccggcctggagccacccgataacttgccggtgcccataaacc 6408
 Db 7075 gccgaacagggacctcccaatccagaaca----- 7103
 QY 6409 tgctgcggcctgacataaacacagccgctagacaacagtaacctgagattgattctacc 6468
 Db 7104 -----gagggcatctcttaacccctctgtattacaatacaacctctgatacgtcccaact 7158
 QY 6469 aacacgcctagaaactcccccaggtgtctgtttaagacagagacagagactctcagctc 6528
 Db 7159 ga-----gcctaaactcaactataaacacaggggcgaactttttaacctc 7203
 QY 6529 atccaggagctttccaagccatcaactcaacccagacccctgatacgaacttcttctgttg 6588
 Db 7204 atccaggagcttttcaagctcttaactccacgactccagaggtactcttctgttg 7263
 QY 6589 ctgtctctctcagggcctctcttattatagagggatggctaaagaaagaaattcaat 6648
 Db 7264 ctgtctctagctcggggccaccttactatgagggatggctagaggggaaattcaat 7323
 QY 6649 gtgacaaagagcatagaaatcaatgtacatgggggtcc 6687
 Db 7324 gtgacaaagaaacatagagacaaatgtacatggggatcc 7362

RESULT 7

AAAF77726
 ID AAF77726 standard; DNA; 7333 BP.
 XX AC AAF77726;
 XX DF 23-MAY-2001 (first entry)
 XX DE Defective retroviral genome isolated from PK-15 cell line.
 XX KW Retrovirus; graft transplantation; xenotransplantation; PK-15 cell line;
 XX SS ss.
 OS Unidentified.
 XX PN US6190861-B1.
 XX PD 20-FEB-2001.
 XX PF 13-DEC-1996; 96US-0766528.
 XX PR 14-DEC-1995; 95US-0572645.
 XX PA (GEMO) GEN HOSPITAL CORP.
 XX PI Fishman JA;
 XX DR WPI; 2001-256211/26.
 DR P-PSDB; AAB73282, AAB73283, AAB73284.

XX

PT

PT

PT

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ

Sequence 7333 BP; 1984 A; 1788 C; 1885 G; 1676 T; 0 other;

Query Match

Best Local Similarity 66.9%; Score 5489.8; DB 22; Length 7333;
 Matches 6856; Conservative 0; Mismatches 402; Indels 1005; Gaps 12;

QY 1 gtggtgtacagactgtggccccagcgcgttggaaataaaatctctgtgtgttcac 60

Db 11 gtggtgtacagactgtggccccagcgcgttggaaataaaatctctgtgttcac 70

QY 61 aagacgcctctcgtgagtgattgggtgtgcctctctccagccgacgaggggat 120

Db 71 aagacgcctctcgtgagtgattgggtgtgcctctctccagccgacgaggggat 130

QY 121 tgtcttttactgccccttcatttgggtgtgcctctctccagccgacgaggggat 180

Db 131 tgtcttttactgccccttcatttgggtgtgcctctctccagccgacgaggggat 190

QY 181 caccgcagaaacacgacttgaggtaaaggatccccctttggaacgtgtgtgtgcgc 240

Db 191 caccgcagaaacacgacttgaggtaaaggatccccctttggaacgtgtgtgtgcgc 250

QY 241 ggcgtctctgtctgtgagttctgttttgcgtgagcgttttcgtgtgtgtgtgtgt 300

Db 251 ggcgtctctgtctgtgagttctgttttgcgtgagcgttttcgtgtgtgtgtgtgt 310

QY 301 ctacagaccgtaagacgacttgaggactgtgatcagcagacgtgtgtgtgtgtgtgtgt 360

Db 311 ctacagaccgtaagacgacttgaggactgtgatcagcagacgtgtgtgtgtgtgtgt 370

QY 361 ccaccttggggagc 420

Db 371 ccaccttggggagc 430

QY 421 gtccgtcagagagacccaggttctgttttgaagcgaagcttccccctccgcgcgcgcgc 480

Db 431 gtccgtcagagagacccaggttctgttttgaagcgaagcttccccctccgcgcgcgcgc 490

QY 481 acctctttgcctgcttctgtggaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 540

Db 491 acctctttgcctgcttctgtggaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 550

QY 541 ctgtctcgtgtgtcttttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 600

Db 551 ctgtctcgtgtgtcttttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 610

QY 601 tgactaccccccttagtttgactctcgcacattggactgaagttagatccagggctcata 660

Db 611 tgactaccccccttagtttgactctcgcacattggactgaagttagatccagggctcata 670

QY 661 atttgtcagttcaggttaagaggggaccttggcagacttctcgtcctctgaatggccaa 720

Db 671 atttgtcagttcaggttaagaggggaccttggcagacttctcgtcctctgaatggccaa 730

Assessing risk of endogenous retroviruses in clinical practice and in xenotransplantation, comprises using probe sequences derived from swine or miniature swine retroviral genome.

Claim 1; Fig 2; 127pp; English.

The present invention relates to a method for screening a cell or tissue for the presence or expression of a retrovirus (RV), comprising contacting a target nucleic acid from the cell or tissue with a second nucleic acid from the present invention (e.g. the present sequence or a fragment thereof). The method is useful for RV detection and to assess graft transplantation risk. Screening of animals allows the elimination of donors with active replication of known viruses. Inactive proviruses can be detected and inactivated, allowing identification and elimination of potential human pathogens derived from swine in a manner not possible in the outbred human organ donor population and is important to the development of human xenotransplantation.

QY 721 cattcgatgttgatggccatcagagggaacctttaattctgaattatctcgtgcta 780
DB 731 cattcgatgttgatggccatcagagggaacctttaattctgaattatctcgtgcta 790
QY 781 aggcaatcatctttcagactggaccggtctcatctgatacaggagccctatactta 840
DB 791 aggcaatcatctttcagactggaccggtctcatctgatacaggagccctatactta 850
QY 841 cgtggcaagatttgagcaagatcctccgcatgggttaaacctgggttaataaaccaa 900
DB 851 cgtggcaagatttgagcaagatcctccgcatgggttaaacctgggttaataaaccaa 910
QY 901 gaaagccaggttcccggaatctcgtcttggagagaaaaaacaacactcggccgaaaaag 960
DB 911 gaaagccaggttcccggaatctcgtcttggagagaaaaaacaacactcggccgaaaaag 970
QY 961 tcgagccctctctcgtatatacccgagatcgagagccgagctggccggaaccc 1020
DB 971 tcgagccctc-tcctcgtatatacccgagatcgagagccgagctggccggaaccc 1029
QY 1021 caactgttcccccaaccccttatccagcacaggtgctgtgaggaa-cctctgcccct 1079
DB 1030 caactgttcccccaaccccttatccagcacaggtgctgtgaggaaacccctctgcccct 1089
QY 1080 cctggagctccggttggtagggacctgctgccgggactcggagccggagagcgccaccc 1139
DB 1090 cctggagctccggttggtagggacctgctgccgggactcggagccggagagcgccaccc 1149
QY 1140 ccgagcgagacagacagatcgatataccgctcgcaactatggccctcccatgcca 1199
DB 1150 ccgagcgagacagacagatcgatataccgctcgcaactatggccctcccatgcca 1209
QY 1200 gggcgcaattgagccctccatattggccctttctctcgcagatctctataattgg 1259
DB 1210 gggcgcaattgagccctccatattggccctttctctcgcagatctctataattgg 1269
QY 1260 aaaaactaacctccccctttctggagagatccccaacgctcacggggttggtagctcc 1319
DB 1270 aaaaactaacctccccctttctggagagatccccaacgctcacggggttggtagctcc 1329
QY 1320 ctatgttctctacacgctacttggagatgtgtcaacagctgctgcagacactcttc 1379
DB 1330 ctatgttctctacacgctacttggagatgtgtcaacagctgctgcagacactcttc 1389
QY 1380 acaaccagagcagagagaaattctgttagaggtcagaaaaaatgttctcgggcccag 1439
DB 1390 acaaccagagcagagagaaattctgttagaggtcagaaaaaatgttctcgggcccag 1449
QY 1440 gggcgacccacgaggttgcaaaatgagatgacatggatttcccttgactcggcccgt 1499
DB 1450 gggcgacccacgaggttgcaaaatgagatgacatggatttcccttgactcggcccgt 1509
QY 1500 tgggactacaacacgctgaaggttagggagcttgaaaaatctatccaggtctcgtg 1559
DB 1510 tgggactacaacacgctgaaggttagggagcttgaaaaatctatccaggtctcgtg 1569
QY 1560 gcgggttccgggcccctcaaacgcccactaatgttgctaaaggttaagagaggtatg 1619
DB 1570 gcgggttccgggcccctcaaacgcccactaatgttgctaaaggttaagagaggtatg 1629
QY 1620 cagggacccaacacccctcctcgtatttctgagaggtcatggaagccttcaggcgg 1679
DB 1630 cagggacccaacacccctcctcgtatttctgagaggtcatggaagccttcaggcgg 1689
QY 1680 ttcacccctttgatactacagagcccaagaaacccatcagtgccctggcctcatt 1739
DB 1690 ttcacccctttgatactacagagcccaagaaacccatcagtgccctggcctcatt 1749
QY 1740 gggcagtcgctctgatatcaggaagaaacttcagagactggaaggtttacaggaggt 1799
DB 1750 gggcagtcgctctgatatcaggaagaaacttcagagactggaaggtttacaggaggt 1809

QY 1800 gagttaactgtatctagtgagagaggcagagaagtggtattacagaaggagacagaag 1859
DB 1810 gagttaactgtatctagtgagagaggcagagaagtggtattacagaaggagacagaag 1869
QY 1860 gagaaggaacagagaaaaaaggagagagaagaaggaggaagacagctgatacagg 1919
DB 1870 gagaaggaacagagaaaaaaggagagagaagaaggaggaagacagctgatacagg 1929
QY 1920 cagagagaagatttgactaagatcttggccagctggttgaagggaagacagcaggagg 1979
DB 1930 caagagaagaatttgactaagatcttggccagctggttgaagggaagacagcaggagg 1989
QY 1980 agagagagagattttaggaaaatttagtcaggccctagacagtcagggaacctgggcaat 2039
DB 1990 agagagagagattttaggaaaatttagtcaggccctagacagtcagggaacctgggcaat 2049
QY 2040 agagcccaactcagaaaggaccagctgtgcgtattgttaagaaaaagacactgggcaagg 2099
DB 2050 agagcccaactcagaaaggaccagctgtgcgtattgttaagaaaaagacactgggcaagg 2109
QY 2100 aactgccccaaagaaagaaacaaagacagaggtcctagctctagaagaataaagat 2159
DB 2110 aactgccccaaagaaagaaacaaagacagaggtcctagctctagaagaataaagat 2169
QY 2160 tagggagagcgggttcggacccctccccgagccaggggttaactttgaaggtggagggg 2219
DB 2170 tagggagagcgggttcggacccctccccgagccaggggttaactttgaaggtggagggg 2229
QY 2220 caaccagttgagttcctcgttgataccggagggagcatttcagtgctgtacaacatt 2279
DB 2230 caaccagttgagttcctcgttgataccggagggagcatttcagtgctgtacaacatt 2289
QY 2280 ggaataactaaagaaaaaaatcctcgttgatgggtgcacaggggcaacggcagtatcca 2339
DB 2290 ggaataactaaagaaaaaaatcctcgttgatgggtgcacaggggcaacggcagtatcca 2349
QY 2340 tggactacccgaagaacccgttgacttgggagtgggagcgggttaaccactcgtttcgtgctc 2399
DB 2350 tggactacccgaagaacccgttgacttgggagtgggagcgggttaaccactcgtttcgtgctc 2409
QY 2400 atccctgagtgccagtlacccctctctagtagagacttactgaccaagatggaggtctca 2459
DB 2410 atccctgagtgccagtlacccctctctagtagagacttactgaccaagatggaggtctca 2469
QY 2460 atttctttgaaacagaagacagaaagtgtcgtgaaataacaaacccactcactgtgtg 2519
DB 2470 atttctttgaaacagaagacagaaagtgtcgtgaaataacaaacccactcactgtgtg 2529
QY 2520 accctccaatttagatgataatcagactatattctccccagtgaaagcctgatacaagat 2579
DB 2530 accctccaatttagatgataatcagactatattctccccagtgaaagcctgatacaagat 2589
QY 2580 atacagtcctggttgagcagtttccccagcctggcgagaaacggcaggtgggtttg 2639
DB 2590 atacagtcctggttgagcagtttccccagcctggcgagaaacggcaggtgggtttg 2649
QY 2640 gcaaaacaggttccccccacaggttattcaactgaaggccagtgctacacagatcagctc 2699
DB 2650 gcaaaacaggttccccccacaggttattcaactgaaggccagtgctacacagatcagctc 2709
QY 2700 agacagtaaccttggagtagagaggtcagagagaaatttggccatgttcaaaagatta 2759
DB 2710 agacagtaaccttggagtagagaggtcagagagaaatttggccatgttcaaaagatta 2769
QY 2760 atccaaaggggcatcctagttcctgtccaaatcccccttggaatactccccctgtaacgggt 2819
DB 2770 atccaaaggggcatcctagttcctgtccaaatcccccttggaatactccccctgtaacgggt 2829
QY 2820 aggaacactgggacaatgattatgacagcagttacaggagcttgagagaggtcaataaaagg 2879
DB 2830 aggaacactgggacaatgattatgacagcagttacaggagcttgagagaggtcaataaaagg 2889
QY 2880 gtgcaggacatacaccccaacgggtcccgaaaccttataacctctgtgagcgcctcccgct 2939

Db 2890 |||||gtgcaggacataccaccacggtcccgaaacctctataaacctcttgagcgccctcccgct 2949
QY 2940 gaacggaactggtacacagtttgacttaaaagatgcttctctctgctgagattacac 2999
Db 2950 gaacggaactggtacacagtttgacttaaaagatgcttctctctgctgagattacac 3009
QY 3000 cccactagccaaaccttttgccttcgaattggagagatccagggtacggggaagaccggg 3059
Db 3010 cccactagccaaaccttttgccttcgaattggagagatccagggtacggggaagaccggg 3069
QY 3060 cagctcaactggaccgcgactgcccgaagggttcaagaactcccgaacctatttgacgaa 3119
Db 3070 cagctcaactggaccgcgactgcccgaagggttcaagaactcccgaacctatttgacgaa 3129
QY 3120 gccctacacaggacctgccaacttcaggtacggtacccacacctcaggtgacctctccag 3179
Db 3130 gccctacacaggacctgccaacttcaggtacggtacccacacctcaggtgacctctccag 3189
QY 3180 tacgtggatgacctgcttctgcgggagcccaaacaggagctgcttagaagggtacgaag 3239
Db 3190 tacgtggatgacctgcttctgcgggagcccaaacaggagctgcttagaagggtacgaag 3249
QY 3240 gaactactgctgaattgctgaactgagctacagagcctctgcttaagaaggccagatt 3299
Db 3250 gaactactgctgaattgctgaactgagctacagagcctctgcttaagaaggccagatt 3309
QY 3300 tqcaggagaggttaacatactgggttacagtttgcggggtcgagcgatgctgacg 3359
Db 3310 tqcaggagaggttaacatactgggttacagtttgcggggtcgagcgatgctgacg 3369
QY 3360 gaggccaggaagaaacttgtagtcagataccggcccaacacagcccaacaaagtgaga 3419
Db 3370 gaggccaggaagaaacttgtagtcagataccggcccaacacagcccaacaaagtgaga 3429
QY 3420 gaggttttgggaagctggatttgcagactgtgatcccggttgcgaccttagca 3479
Db 3430 gaggttttgggaagctggatttgcagactgtgatcccggttgcgaccttagca 3489
QY 3480 gccccacttaccgcttaacaaagaaaggggattctcctgggtcctgagcaccag 3539
Db 3490 gccccacttaccgcttaacaaagaaaggggattctcctgggtcctgagcaccag 3516
QY 3540 aaggcatttgatgctatacaaaaggccctgtgagcgcaactgtctgcccctccctgac 3599
Db 3517 ----- 3516
QY 3600 gtaactaaacctttaccctttatgtggtgagcgtaaggagtagcccgagggtttta 3659
Db 3517 ----- 3516
QY 3660 acccaaaccttaggaccatggaggagacctgttgctacctgtcaaaagagcttgatcct 3719
Db 3517 ----- 3516
QY 3720 gtacgcagtggttggcccgatgtctgaagctatcgagctgtgcccactactggtcaag 3779
Db 3517 ----- 3516
QY 3780 gacgtgacaaatgactttgggacagataataactgtaataagccccccatgcatggag 3839
Db 3517 ----- 3516
QY 3840 aacatcgttcggcagccccccagacggtggtgacccaacgcccgcgcatgacccactatcaa 3899
Db 3517 ----- 3516
QY 3900 agcctgcttctcacagagagggttcacttctgctccaccagccgctctcaacccctgcaact 3959
Db 3517 ----- 3516
QY 3960 cttctgctgaagagactgatgaaccagtgactcatgattgattgattgag 4019

Db 3517 ----- 3516
QY 4020 gagactgggtccgcgaaggaccttaacagacataccgctgactggagaagtgtcaacctgg 4079
Db 3517 ----- 3516
QY 4080 ttcactgaocggaagcagctatgtggtggaaggttaagagatggtggtggcgagtcggtg 4139
Db 3517 ----- 3516
QY 4140 gacgggaccgcgacgatcttgccgacgacctgcccgaaggaacttcagcgcaaaagct 4199
Db 3517 ----- 3516
QY 4200 gagctcatggccctcagcgaagcttggcggtggccggaagggaatccataaaacattat 4259
Db 3517 ----- 3516
QY 4260 acggacagcagggtatgcttgcgactgcacgtcacggtgcacgggccaactataaaacagg 4319
Db 3517 -----aaag 3521
QY 4320 ggttgcctaccctcagcagggaggaataaaacaaagagaaattcttaagcctatta 4379
Db 3522 ggttgcctaccctcagcagggaggaataaaacaaagagaaattcttaagcctatta 3581
QY 4380 gaagccttacatttgcacaaaggctagctattatatacctgtcctggacatcagaagcc 4439
Db 3582 gaagccttacatttgcacaaaggctagctattatatacctgtcctggacatcagaagcc 3641
QY 4440 aaagatctcatctagaggaacccagatggtgacctgggttgcgaagcagcagccag 4499
Db 3642 aaagatctcatctagaggaacccagatggtgacctgggttgcgaagcagcagccag 3701
QY 4500 gctgttaacctctgcctataatagaaaccccaagcccaagcccaagacacagtagac 4559
Db 3702 gctgttaacctctgcctataatagaaaccccaagcccaagacacagtagac 3761
QY 4560 acctagaagactggcgaagagataaaagacagtagacagttctctgagactccggaggg 4619
Db 3762 acctagaagactggcgaagagataaaagacagtagacagttctctgagactccggaggg 3821
QY 4620 acctgtataacctcaatgggaaggaataacctcccccaaaaggggttagaatgtc 4679
Db 3822 acctgtataacctcaatgggaaggaataacctcccccaaaaggggttagaatgtc 3881
QY 4680 caacagatacatcgtctaaaccccttaggaactaaacacctgcagcagttggtcagaaca 4739
Db 3882 caacagatacatcgtctaaaccccttaggaactaaacacctgcagcagttggtcagaaca 3941
QY 4740 tcccttatcatgttctgaggtaccagggagtggtgactcgtggtcaaacatttgtg 4799
Db 3942 tcccttatcatgttctgaggtaccagggagtggtgactcgtggtcaaacatttgtg 4001
QY 4800 cctgtcagctggttaatgctaatccttcagataacctccagaaagagactaaagggga 4859
Db 4002 cctgtcagctggttaatgctaatccttcagataacctccagaaagagactaaagggga 4061
QY 4860 agccaccagcgtcactggagtggttcaactgaggtaaagccgcttaaatcagga 4919
Db 4062 agccaccagcgtcactggagtggttcaactgaggtaaagccgcttaaatcagga 4121
QY 4920 acaaaatctattgtgttttttagacaccttttcaggatgggtgagagcttactact 4979
Db 4122 acaaaatctattgtgttttttagacaccttttcaggatgggtgagagcttactact 4181
QY 4980 aagaaaagagacttcaaccggtgggttaagaaataactggagaaatttttccaaattt 5039
Db 4182 aaaaaagagacttcaaccggtgggttaagaaataactggagaaatttttccaaattt 4241
QY 5040 ggaatacctaaggttaataggttcagacaaatggtccagcttctggtccaggttaagtcag 5099
Db 4242 ggaatacctaaggttaataggttcagacaaatggtccagcttctggtccaggttaagtcag 4301

Db 6331 tggttctacagacagaaggggttagatctgttatcttctaaagagagggttatgtg 6390
QY 7256 cagccttaaaagagaatgttgcctctctgttagatcactcagggccatcagagactcca 7315
Db 6391 tagccttaaaagagaatgttgcctctctgttagatcactcagggccatcagagactcca 6450
QY 7316 tgaacaagcttagaaaaagttagagagcctcgaagggaagagagagagagagaggggt 7375
Db 6451 tgaacaagcttagaaaaagttagagagcctcgaagggaagagagagagagagaggggt 6510
QY 7376 ggttgaaggatggttcaacaggtctccttgatgagacacccctcttctgctcagcg 7435
Db 6511 ggttgaaggatggttcaacaggtctccttgatgagacacccctcttctgctcagcg 6570
QY 7436 ggcctctagtagtctcctctgtaactacagttggccttgcttaataataggttg 7495
Db 6571 ggcctctagtagtctcctctgtaactacagttggccttgcttaataataggttg 6630
QY 7496 ttgctctttagagaaagagtagtcagtcacagatcattgtaacttaggcaacagtcac 7555
Db 6631 ttgctctttagagaaagagtagtcagtcacagatcattgtaacttaggcaacagtcac 6690
QY 7556 aaggccttctgagccaaagagaactgacctctagccttccagttctaaagattagaaact 7615
Db 6691 aaggccttctgagccaaagagaactgacctctagccttccagttctaaagattagaaact 6750
QY 7616 attaaagacagaagtgagggaatgaagagatgaataatgaacactaacctccccaagac 7675
Db 6751 attaaagacagaagtgagggaatgaagagatgaataatgaacactaacctccccaagac 6810
QY 7676 ccaggaagttaataaagctcctaattgcccgaaattacagaccctgctggctgcaggt 7735
Db 6811 ccaggaagttaataaagctcctaattgcccgaaattacagaccctgctggctgcaggt 6870
QY 7736 aataggttagaggtcacactctctattgttccagggcctgctctatctgcctaaagtaag 7795
Db 6871 aataggttagaggtcacactctctattgttccagggcctgctctatctgcctaaagtaag 6930
QY 7796 ataacaggaatgaggtgactaatcgtcttctgtgattctgtataaactgactggccact 7855
Db 6931 ataacaggaatgaggtgactaatcgtcttctgtgattctgtataaactgactggccact 6990
QY 7856 agaagaatttaataacatgacagccctagtagcctatctcactcactcactcactcactc 7915
Db 6991 agaagaatttaataacatgacagccctagtagcctatctcactcactcactcactcactc 7050
QY 7916 tgccaggagcccgacagatgcggacccctcgagctattttaaatgattggtccacgg 7975
Db 7051 tgccaggagcccgacagatgcggacccctcgagctattttaaatgattggtccacgg 7110
QY 7976 agcggggtctcgatattttaaatgattggttcacagcgcggtggtcgtcgtcgtatattt 8035
Db 7111 agcggggtctcgatattttaaatgattggttcacagcgcggtggtcgtcgtcgtatattt 7170
QY 8036 aaaatgattggttgcacacagccttgcgtgaccccccataaaagctgccgatt 8095
Db 7171 aaaatgattggttgcacacagccttgcgtgaccccccataaaagctgccgatt 7230
QY 8096 ccgactcgggcccagcagctctcaccctcgctggtgtagcactgtgggcccagcgcg 8155
Db 7231 ccgactcgggcccagcagctctcaccctcgctggtgtagcactgtgggcccagcgcg 7290
QY 8156 cttgggaataaaatcctctgctgtttgcatcaataaaaaa 8198
Db 7291 cttgggaataaaatcctctgctgtttgcatcaataaaaaa 7333

RESULT 8

AAT74883

ID AAT74883 standard; cDNA: 7393 BP.

XX

AC AAT74883;

XX DT 09-FEB-1998 (first entry)
XX DE Porcine retrovirus cDNA (defective).
XX KW Retrovirus; porcine; GAG protein; POL protein; ENV protein;
KW xenotransplantation; infectious; provirus; organ transplant; donor;
KW activated virus; PCR; ss.
XX OS Porcine retrovirus.
XX FH Key Location/Qualifiers
FT CDS 598..2172
FT /tag= a
FT mat_peptide 598..2169
FT /tag= b
FT /note= "putative GAG protein"
FT CDS 2320..4737
FT /tag= c
FT /note= "putative POL coding region (partial) as
FT described in the specification"
FT mat_peptide 2320..3522
FT /tag= d
FT /note= "putative POL protein (partial)"
FT mat_peptide 3516..4328
FT /tag= e
FT /note= "putative POL protein (partial)"
FT mat_peptide 4332..4748
FT /tag= f
FT CDS 4738..6725
FT /tag= g
FT /note= "putative ENV coding region (partial) as
FT described in the specification"
FT mat_peptide 4752..6722
FT /tag= h
FT /note= "ENV protein (partial)"
XX PN WO9721836-A1.
XX PD 19-JUN-1997.
XX PF 13-DEC-1996; 96WO-US19680.
XX PR 14-DEC-1995; 95US-0572645.
XX PA (GEO) GEN HOSPITAL CORP.
XX PI Fishman JA;
XX DR WPI: 1997-332804/30.
DR P-PSDB; AAW32091-W32095.
XX CC New nucleic acid from porcine retroviruses - used for detecting
XX PT viruses in transplant or other tissue and for assessing risk of
XX PS transmitting infection to graft recipient
XX CC Claim 16; Fig 2; 128pp; English.
XX CC This cDNA sequence represents a defective purified swine retrovirus
CC found in PK-15 cells containing the putative coding regions for viral
CC GAG, POL and ENV proteins. There are a few in frame stop codons and
CC apparent frame shifts in the given coding sequence which alter features
CC of the translation. This sequence and PCR fragments generated from the
CC sequence (see AAT74812-74882) could be used to screen organs for the
CC presence of porcine retroviruses prior to xenotransplantation.
CC Transplantation can increase the likelihood of retroviral activation if
CC intact and infectious proviruses are present. The porcine retroviral
CC sequence can be used to generate probes to determine the level (e.g.
CC copy number) of intact (i.e. potentially replicating) porcine provirus
CC sequences in a strain of xenograft transplantation donors. It can be
CC used to detect mutations, genetic lesions or viral recombinants and to
CC determine the histological localisation of activated retrovirus. Using
CC Polymerase Chain Reaction DNA Quantitation (PQO) on blood mononuclear

CC cells, infectivity titration and susceptibility testing can be
CC performed. Ultimately animal donors without intact porcine retroviral
CC sequences or a lower copy number of viral elements could be selected.
XX
SQ Sequence 7393 BP; 2004 A; 1801 C; 1896 G; 1692 T; 0 other;

Query Match 66.0%; Score 5416.6; DB 18; Length 7393;
Best Local Similarity 82.4%; Pred. NO. 0;
Matches 6854; Conservative 0; Mismatches 404; Indels 1065; Gaps 13;

QY 1 gtggtgtacgactgtggccccagcgcttggaataaaaaatcctctgtgttgcate 60
DB 11 gtggtgtacgactgtggccccagcgcttggaataaaaaatcctctgtgttgcate 70
QY 61 aagacgccttcgtgaagtattgggtgtcgccctctccgagccgagcagaggggat 120
DB 71 aagacgccttcgtgaagtattgggtgtcgccctctccgagccgagcagaggggat 130
QY 121 tgtctttactggcctttcatttgggtgtggcggaatacctcgacacccctta 180
DB 131 tgtctttactggcctttcatttgggtgtggcggaatacctcgacacccctta 190
QY 181 caccagaaacgacttgaggtaaaggatccctttggaacgtgtgtgtcgcc 240
DB 191 caccagaaacgacttgaggtaaaggatccctttggaacatgtgtgtcgcc 250
QY 241 ggcgtctctgttcgactgtctgttttgcgtgatgcgcctttcggttgcagctct 300
DB 251 ggcgtctctgttcgactgtctgttttgcgtgatgcgcctttcggttgcagctct 310
QY 301 ctgagaccgttaagactggagactgtatcagcagacgtgtcagagatcacagctg 360
DB 311 ctgagaccgttaagactggagactgtatcagcagacgtgtcagagatcacagctg 370
QY 361 caacctgtgggagccccggagggtggggagagcagggagcctgtgtgtctctact 420
DB 371 caacctgtgggagccccggagggtggggagagcagggagcctgtgtgtctctact 430
QY 421 gtcggtcagagaccgagttctgttgaagcgcgacgggtcgctgtgtcgtgtgtgttt 540
DB 431 gtcggtcagagaccgagttctgttgaagcgcgacgggtcgctgtgtcgtgtgttt 550
QY 481 actcttttgcgtctgttgaagcgcgacgggtcgctgtgtcgtgtgtgtgttt 540
DB 491 actcttttgcgtctgttgaagcgcgacgggtcgctgtgtcgtgtgtgtgttt 550
QY 541 ctgtctcgtgtctgttctgtcgtctgtctgtctgtctgtctgtctgtctgtctgt 600
DB 551 ctgtctcgtgtctgttctgtcgtctgtctgtctgtctgtctgtctgtctgtctgt 610
QY 601 tgactaccccccttagttgactctcgaccattgagactgaagttagatccagggctcata 660
DB 611 tgactaccccccttagttgactctcgaccattgagactgaagttagatccagggctcata 670
QY 661 attgtcagttcaggttaagaaggacattggcagacttctgtcctctgaatggccaa 720
DB 671 attgtcagttcaggttaagaaggacattggcagacttctgtcctctgaatggccaa 730
QY 721 cattcagatttgatggccatcagaggggacattttaaattctgaaattatctcgtgttta 780
DB 731 cattcagatttgatggccatcagaggggacattttaaattctgaaattatctcgtgttta 790
QY 781 aggcatactatttcagactgagcccgctctcaatcctctgatcagagccctatactctta 840
DB 791 aggcatactatttcagactgagcccgctctcaatcctctgatcagagccctatactctta 850
QY 841 cgtggcaagatttggcagaagatcctccgcatgggttgaacatcgtgttaataaacaa 900
DB 851 cgtggcaagatttggcagaagatcctccgcatgggttgaacatcgtgttaataaacaa 910
QY 901 gaaagccaggtcccccgaatcctgtcttggagagaaaaaacacactcggccgaaaaag 960
|||||

DB 911 gaaagccaggtcccccgaatcctgtcttggagagaaaaaacacactcggccgaaaaag 970
QY 961 tcgagccctcttctctgtatctaccocgagatcgagggagcgcgacttggccgaaaccc 1020
DB 971 tcgagccctc-tctctgtatctaccocgagatcgagggagcgcgacttggccgaaaccc 1029
QY 1021 caactctgttcccccaccccccttatccagcacaggtgtgtgagggga-cctctgcacct 1079
DB 1030 caactctgttcccccaccccccttatccagcacaggtgtgtgagggacctctgcacct 1089
QY 1080 cctgagctcccggtgtgagggagcctgtccgagactcggagccgagagggcgccacc 1139
DB 1090 cctgagctcccggtgtgagggagcctgtccgagactcggagccgagagggcgccacc 1149
QY 1140 ccggagcgagacagacagatcgcgatatcaccgtcgcacctatggccctcccatgcc 1199
DB 1150 ccggagcgagacagacagatcgcgatatcaccgtcgcacctatggccctcccatgcc 1209
QY 1200 gggggccaattgcagccctccagtttggcccttttctctcagatctctataattgg 1259
DB 1210 gggggccaattgcagccctccagtttggcccttttctctcagatctctataattgg 1269
QY 1260 aaaaataaccatccccctttctcgagatcccccaacgctcaccgggttggtagtcc 1319
DB 1270 aaaaataaccatccccctttctcgagatcccccaacgctcaccgggttggtagtcc 1329
QY 1320 ctatgttctctcacccacttcttggatgttgcacagctgtcagacactcttc 1379
DB 1330 ctatgttctctcacccacttcttggatgttgcacagctgtcagacactcttc 1389
QY 1380 acaaccgagcgagacagaaattctgttagaggtgagaaaaatgttctggggcgac 1439
DB 1390 acaaccgagcgagagagaaattctgttagaggtgagaaaaatgttctggggcgac 1449
QY 1440 gggcgaccacgcagctgtcgaataagattgacatgggttctcttgactcgcgcgg 1499
DB 1450 gggcgaccacgcagctgtcgaataagattgacatgggttctcttgactcgcgcgg 1509
QY 1500 tgggactacaacagcgctgaaggtagggagagcttgaatactatcgccaggtctgg 1559
DB 1510 tgggactacaacagcgctgaaggtagggagagcttgaatactatcgccaggtctgg 1569
QY 1560 gcgggtctccggggcgccctcaagcgcccaactaatttggctaaggtaagagaggtgat 1619
DB 1570 gcgggtctccggggcgccctcaagcgcccaactaatttggctaaggtaagagaggtgat 1629
QY 1620 caggagccgaacgaacctccctcggtatttcttgagaggtctatggagccttcaggcgg 1679
DB 1630 caggagccgaacgaacctccctcggtatttcttgagaggtctatggagccttcaggcgg 1689
QY 1680 ttcaccccttttgatctctacatcagagggccagaaagcctcagtgcccttgcatt 1739
DB 1690 ttcaccccttttgatctctacatcagagggccagaaagcctcagtgcccttgcatt 1749
QY 1740 gggcagtcgctctggatatcaggaagaaacttcagagactggaggttaccagaggtct 1799
DB 1750 gggcagtcgctctggatatcaggaagaaacttcagagactggaggttaccagaggtct 1809
QY 1800 gagtacgtgtatctagttagagagcgagagaggtgtattacagaagggagacagagag 1859
DB 1810 gagtacgtgtatctagttagagagcgagagaggtgtattacagaagggagacagagag 1869
QY 1860 gagaaggaacagagaaaaagagagagagagagagagagagagagagagagagag 1919
DB 1870 gagaaggaacagagaaaaagagagagagagagagagagagagagagagagagag 1929
QY 1920 caagagaagaatttgcataagatcttggccagctgttgaaggaagagcagcagagag 1979
DB 1930 caagagaagaatttgcataagatcttggccagctgttgaaggaagagcagcagagag 1989
QY 1980 agagagagagattttagaaaaattaggttcagggcctagacagtcagggaaacttgggcaat 2039
DB 1990 agagagagagattttagaaaaattaggttcagggcctagacagtcagggaaacttgggcaat 2049

Db 3517 ----- 3516
QY 4260 acggacagcaggtatgcttgcgactgcacacgtacacgggcccacatctataaaacaaagg 4319
Db 3517 -----aaagg 3521
QY 4320 ggggtgttacctcagcagggaggaaataaaagaaacaaagagaaattctaagcctatta 4379
Db 3522 ggggtgttacctcagcagggaggaaataaaagaaacaaagagaaattctaagcctatta 3581
QY 4380 gaagccttacattgccccaaaaaggctagctattatcacctgtcctggacatcagaagacc 4439
Db 3582 gaagccttacattgccccaaaaaggctagctattatcacctgtcctggacatcagaagacc 3641
QY 4440 aaagatctcatatctagagggaaccagatggctgaccgggttgccaagcaggacccacag 4499
Db 3642 aaagatctcatatctagagggaaccagatggctgaccgggttgccaagcaggacccacag 3701
QY 4500 gctgttaaccttctgctataataatagaacgcgccaaagccccaagccccaagacccagacagtac 3761
Db 3702 gctgttaaccttctgctataataatagaacgcgccaaagccccaagacccccaagacagtac 3761
QY 4560 accctagaagactggcagagagataaaagatagaccgggttctctgagactccggagggg 4619
Db 3762 accctagaagactggcagagagataaaagatagaccgggttctctgagactccggagggg 3821
QY 4620 acctgctatacctcatatgggaaggaaatcctgcccccaagaaagaggttagaataatctc 4679
Db 3822 acctgctatacctcatatgggaaggaaatcctgcccccaagaaagaggttagaataatctc 3881
QY 4680 caacagatatactgcttaacccacctaggaactaaacacctgcagcagttggtcagaaca 4739
Db 3882 caacagatatactgcttaacccacctaggaactaaacacctgcagcagttggtcagaaca 3941
QY 4740 tccccctatctgcttctgagctaccaggagtggtgctgactcgggtggtcaaacattgtgtg 4799
Db 3942 tccccctatctgcttctgagctaccaggagtggtgctgactcgggtggtcaaacattgtgtg 4001
QY 4800 cccctgacgtggttaactgcttaactctccagataactccagaaagagactaaggga 4859
Db 4002 cccctgacgtggttaactgcttaactctccagataactccagaaagagactaaggga 4061
QY 4860 agccacccagcgtcactgggaagtggacttcaactgaggttaagccggctaaatacagga 4919
Db 4062 agccacccagcgtcactgggaagtggacttcaactgaggttaagccggctaaatacagga 4121
QY 4920 acaaaatctattggttttctgagacaccttttcaggatgggttagaggcttaacctact 4979
Db 4122 acaaaatctattggttttctgagacaccttttcaggatgggttagaggcttaacctact 4181
QY 4980 aagaaagagactcaacccgtggtggctaaagaaataactggaggaatttttccagattt 5039
Db 4182 aaaaaagagactcaacccgtggtggctaaagaaataactggaggaatttttccagattt 4241
QY 5040 ggaataccttaaggtaaggctcagacaatggtccagcttctcgttccaggttaagtcaag 5099
Db 4242 ggaataccttaaggtaaggctcagacaatggtccagcttctcgttccaggttaagtcaag 4301
QY 5100 ggaactggccaagataattggggataggattggaactgcatgtgcatcacagcccccaagc 5159
Db 4302 ggaactggccaagataattggggataggattggaactgcatgtgcatcacagcccccaagc 4361
QY 5160 tcaggacaggttagagagatgaatagaacccattaaagaccccttaaccaattgaaccaca 5219
Db 4362 tcaggacaggttagagagatgaatagaacccattaaagaccccttaaccaattgaaccaca 4421
QY 5220 gagactggcattaatgattgagtggtctcctgccccttctgctttttaggggtgaggagac 5279
Db 4422 gagactggcattaatgattgagtggtctcctgccccttctgctttttaggggtgaggagac 4481
QY 5280 acccctggacagtttggctgacccctatgaattgctctacgggggacccccccctgtg 5339
Db 5280 acccctggacagtttggctgacccctatgaattgctctacgggggacccccccctgtg 5339

Db 4482 accctggacagtttgggtgacccccctataaaattgctctacgggggacccccccctgtg 4541
QY 5340 gcagaaattgcttgcacatagtgctgctgctgtgtgttccagcctttgttctctaggg 5399
Db 4542 gcagaaattgcttgcacatagtgctgctgctgtgtgttccagcctttgttctctaggg 4601
QY 5400 ctcaaggcgtcgcagtggtgagcagcagcgtggaagcagctccggaggcctactca 5459
Db 4602 ctcaaggcgtcgcagtggtgagcagcagcgtggaagcagctccggaggcctactca 4661
QY 5460 ggaggagacttgcgaagtccacatcgcttccagttggagattgagattctatgttagacgc 5519
Db 4662 ggaggagacttgcgaagtccacatcgcttccagttggagattgagattctatgttagacgc 4721
QY 5520 caccgtgcagaaacctgcagactcgtggaagggacacttatctcgtactttttagaccaca 5579
Db 4722 caccgtgcagaaacctgcagactcgtggaagggacacttatctcgtactttttagaccaca 4781
QY 5580 ccaacggctgtgaaagtcgaaggaaatccccacctggatcgatccacgttaagccg 5639
Db 4782 ccaacggctgtgaaagtcgaagg ----- 4803
QY 5640 gcgcacccctccgactcgggtggagagccgaaaaagactgagaaatcccccttaagctcgc 5699
Db 4804 -----gaatcccttaagctcgc 4822
QY 5700 ctccatcgctggttctctactctactaacaataactccccagccagtagtaaaacgcttat 5759
Db 4823 ctccatcgctggttctctactctactaacaataactcccccaagttaattggtaaacgcttgc 4882
QY 5760 agcagctgcacaccccatagacatttacccttaccctgctgattattgacctgatac 5819
Db 4883 ggacagcccgaaactcccatataaaccttatctcaccctggttacttactgacctggatc 4942
QY 5820 ggggtgctactgtaaatagcactgcagctggttgcctctagagcagcctgggtggcctgaact 5879
Db 4943 aggttaataatataacagcactcaaggaggctcccttgggacccctgggtggcctgaatt 5002
QY 5880 gcatttctgctccgactgatttaaccccgctgta -----aaagcacacccctcccaacct 5933
Db 5003 atatgtctgctcgcagcagtaactccctggtctcaatgaccagggccacaccccgctgtg 5062
QY 5934 agtcctgtagttaggttctctattgctgcccag ---cacagagaaagagaaatctgtgg 5990
Db 5063 actcgtgcttaagggttttcaagttgcccagcagcccccaaaataatgaagaaattgtgg 5122
QY 5991 ggggtctgggaactcctctgttaggagatggagctgcgtcacctcccaagatggagactg 6050
Db 5123 aaatccctcaggatttcttggcaagcaatggagctgcataaacttcaatgaatggggaatgg 5182
QY 6051 gaaatggcagatctctccagcagccgggttaaaattctcctctgtcaattcc ----- 6102
Db 5183 gaaatggcagatctctccagcagccgggttaaaattctcctctgtcaattcc ----- 6102
QY 6103 -----ggcccgggcgaatcaca 6121
Db 5243 ttataatcaatttaattatggccatgggagatggaaagattggcaacagcgggtacaaaa 5302
QY 6122 tgatgaactatataaagataaagactgctcccccatacagacttagattatctaaagataa 6181
Db 5303 agatg-tacgaataagcaataagctgctcaattcgttagacotagattacttaaaaaataa 5361
QY 6182 gtttctcgtg--aaagaaaaacagaaaaatttcaaaagtcgataaaattggtatgagctggg 6239
Db 5362 gtttctcgtgaaaaagaaaaacagaaaaatttcaaaagtcgataaaattggtatgagctggg 5421
QY 6240 gaatagtttttat----tatatggcggggagcaggtccactttaaccattcgcctta 6295
Db 5422 gaatagtgactatgagagctctggagagaaagaaagaaatctgttctgactattcgcctca 5481
QY 6296 ggaatagagcgggagacacccctgtggcaatggagacccataaagactggtgagac 6355
Db 5482 gaatagaaactcagatggacccctcgggttgcataaggaccaaataaagggttggccgaac 5541

PD 30-NOV-2000.
 XX
 PF 24-MAY-2000; 2000WO-US14296.
 XX
 PR 24-MAY-1999; 99US-0135631.
 XX
 PA (MAYO-) MAYO MEDICAL VENTURES.
 XX
 XX Federspiel MJ;
 PI WPI; 2001-032041/04.
 XX
 DR
 XX
 XX Inhibiting or preventing infectious agent transmission in mammalian
 PT transplant recipients, by introducing recombinant DNA comprising DNA
 PT encoding extracellular proteins of the agent into donor cells, such as
 XX swine cells -
 XX
 PS Claim 16; Page 107-109; 144pp; English.
 PS
 CC The present invention provides a method to prevent the transmission of
 CC infectious agents during xenotransplantation. This involves introducing
 CC to donor swine cells a recombinant DNA encoding a peptide fragment from
 CC the infectious agent, and then introducing these cells into the
 CC transplant recipient.
 XX
 SQ Sequence 6076 BP; 1613 A; 1512 C; 1609 G; 1342 T; 0 other;

 Query Match 61.2%; Score 5025.8; DB 22; Length 6076;
 Best Local Similarity 95.7%; Pred. No. 0;
 Matches 5234; Conservative 0; Mismatches 222; Indels 15; Gaps 6;

 QY 1 gtgggtgacgactgtggcccgagcgcttggaataaaaaatcctcttgctgtttgcaac 60
 Db
 601 gfggtgacgactgtggcccgagcgcttggaataaaaaatcctcttgctgtttgcaac 660
 QY 61 aagacgcgtctcgtgagtgatt-tgggggtgcgcctcttcgcagccggagcgaggga 119
 Db
 661 aagacgcgtctcgtgagtgattgaaagggagtcgcttttcgcagccgtggag----- 713
 QY 120 ttgtctttactgccccttatttggctgtggcgggaaatcctcgacacccctt 179
 Db
 714 --ttcttttctgactcttatttgggggctgcctccggatctgtgcggccacccctta 771
 QY 180 acacccgagacccgacttggaggt-aaagggatccctttggaacgtgtgtg--tgtgtc 236
 Db
 772 acacccgagacccgacttggaggtaaaggaatcctcttttaacgtgtatgtatgtacc 831
 QY 237 ggcgggctctgtctgt 296
 Db
 832 ggcgggctctgt 891
 QY 297 tctctcagacgttaagacgtgagacgtgtgatcagcagacgtgtcgttagggagatcacag 356
 Db
 892 tctctcagacgttaagacgtgagacgtgtgatcagcagacgtgtcgttagggagatcacag 951
 QY 357 gctgcacccctggggagcggccggaggtgtgggagagcccgagacgctgtgtgtgtc 416
 Db
 952 gctgcacccctggggagcggccggaggtgtgggagagcccgagacgctgtgtgtgtc 1011
 QY 417 tactctcagacgagacgagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 476
 Db
 1012 tctctcagacgagacgagctgt 1071
 QY 477 tccgactctttgctgt 536
 Db
 1072 tccgactctttgctgt 1131
 QY 537 gtttctgtctgt 596
 Db
 1132 gtttctgtttgt 1191
 QY 597 acagtgactacccccccttagtttgaactctgacccattggactgaagttagatccagggt 656

Db 1192 acggtgacgacccctcttagttgactctcgacattggaactgaagttaataccagggt 1251
 QY 657 cataattgtcagttcagtttaagaaggacacttggcagacatttctgtgcctctgaatgg 716
 Db 1252 cataattgtcagttcagtttaagaaggacacttggcagacatttctgtgtctctgaatgg 1311
 QY 717 ccaacattcgtatttgatggccatcagaggggaccccttaattctgaaattatctcgtgct 776
 Db 1312 cagacattcgtatttgatggccatcagaggggaccccttaattctgaaattatctcgtgct 1371
 QY 777 gtttaaggccaatttttgcagctggacccggctctcattctcagctgagggccctatctc 836
 Db 1372 gtttaaggccaatttttgcagctggacccggctctcattctcagctgagggccctatctc 1431
 QY 837 cttacgtggcagatttggcagaagatcctccgcatcctgggttaaacatcagctgaaataaa 896
 Db 1432 cttacgtggcagatttggcagaagatcctccgcatcctgggttaaacatcctggctgaataa 1491
 QY 897 ccaagaagccaggttcccccgaatcctgctcttgagagaaacacacacactcggccgaa 956
 Db 1492 ccaagaagccaggttcccccgaatcctgctcttgagagaaacacacacactcggctgaa 1551
 QY 957 aaagtgcagccctctctcgttatcccccagatcagagagccgacacttggccgga 1016
 Db 1552 aaagtgcagccctc-tctcatatctaccocgagattgagagccgctgtggccgga 1610
 QY 1017 accccaactctgtcccccaccccttatccagacagaggtgctgtgagggga-cctctgc 1075
 Db 1611 accccaactctgtcccccaccccttatccagagaggtgctgtgaggggaacccctctgc 1670
 QY 1076 cctctctggagcctcgtgtgtgagggacactgctgcgggacactcggagcggagagcgc 1135
 Db 1671 cctctctggagcctcgtgtgtgagggacactgctgcgggacactcggagcggagggcgc 1730
 QY 1136 caccocggagcggacagacgagatcgcgattaccgctgcgcacacttatgcccctccat 1195
 Db 1731 caccocggagcggacagacgagatcgcgacattaccgctgcgcacgctacgcccctccac 1790
 QY 1196 gcaaggggccaattgacgcccctccagatttggcccttttctctgcagatctctataa 1255
 Db 1791 acggggggccaattgacgcccctccagatttggcccttttctctgcagatctctataa 1850
 QY 1256 ttgaaacactaacacatcccttctcggagagatccccacgctcacgggtgtgtgga 1315
 Db 1851 ttgaaacactaacacatcccttctcggagagatccccacgctcacgggtgtgtgga 1910
 QY 1316 gtccttatgttctctcaccagcctacttgggatgattgtcaacagctgctgcagacact 1375
 Db 1911 gtccttatgttctctcaccagcctacttgggatgattgtcaacagctgctgcagacact 1970
 QY 1376 cttcaaacggagggagagagaattctcttagaggtcagaaaaatgttccctggggc 1435
 Db 1971 cttcaaacggagggagagagaattctcttagaggtcagaaaaatgttccctggggc 2030
 QY 1436 cgacggcgacccacacgagttgcaaatgagattgacatgggtttcccttgcactcgc 1495
 Db 2031 cgacggcgacccacacgagttgcaaatgagattgacatgggtttcccttgcactcgc 2090
 QY 1496 cggttgggactcaaacacggtgaaagtagggagacttgaataatctatcccgagctct 1555
 Db 2091 cggttgggactcaaacacggtgaaagtagggagacttgaataatctatcccgagctct 2150
 QY 1556 ggtgaggggtctccggggcgcccaagacggcccaactaatttggctaaagtaagaggt 1615
 Db 2151 ggtgaggggtctccggggcgcccaagacggcccaactaatttggctaaagtaagaggt 2210
 QY 1616 gatcaggggacccgaagacccctcctcgtattcttgagaggtcctatggaagcctcag 1675
 Db 2211 gatcaggggacccgaagacccctcctcgtattcttgagaggtcctatggaagcctcag 2270
 QY 1676 gcggttaccaccccttttgaactcctcagagggcccgagacgctcagtgccctggcct 1735

Db 2271 gcggtacaccccttttgatccacactcagagagcccaaaagccctcagtggtgttggcctt 2330
QY 1736 cattggcgagtcggtcttgatatacagaaagaaacttcagagactgggaaggtttacagga 1795
Db 2331 tataggacagtcagccttgatattagaaagagcttcagagactgggaaggtttacagga 2390
QY 1796 ggcgtgattcagctgactagtgagagagagcagagaagtgattacagaaagggagacaga 1855
Db 2391 ggcgtgattcagctgactagtgagagagagcagagaagtgattacagaaagggagacaga 2450
QY 1856 agagagaaggaagcagagaaag 1915
Db 2451 agagagaaggaagcagagaaagagagagagagagagagagagagagagagagagagag 2510
QY 1916 acggcaagagaagaatttgactaagactctggccgagctggttgaaaggaagagcagcag 1975
Db 2511 acggcaagagaagaatttgactaagactctggccgagctggttgaaaggaagagcagcag 2570
QY 1976 ggaagagagagagagattttagaaatttaggtcagggcccttagacagtcagggaaacctggg 2035
Db 2571 ggaagagagagagagattttagaaatttaggtcagggcccttagacagtcagggaaacctggg 2630
QY 2036 caataggaccctcgcagagagacagctgtgcgtattgttaaagaaagagacactgggc 2095
Db 2631 caataggaccctcgcagagagacacatgtgcattgttaaagaaagagacactgggc 2690
QY 2096 aaggactgcaccaagaggaagaaacaaaggaccgaaggtccttagtaagaagataa 2155
Db 2691 aaggactgcaccaagaggaagaaacaaaggactgaaggtccttagtaagaagataa 2750
QY 2156 agataggaggagacgggttcgagaccctcccgagcccagggtaactttgaaggtgga 2215
Db 2751 agactaggagagacgggttcgagaccctcccgagcccagggtaactttgaaggtgga 2810
QY 2216 ggggcaaccagtgagttcctggttgataccggagcggagagatcagtgctgtacaaac 2275
Db 2811 ggggcaaccagtggttccctggttgataccggagcgaagacattcagtgctactacagc 2870
QY 2276 attagaaactaaagaaagaaataactggtggtgatgggtgccagggcaacggcgagta 2335
Db 2871 attagaaactaaagaaagaaataactggtggtgatgggtgccagggcaacagagta 2390
QY 2336 tccatggactaccgaagacgttgacttgggagtgaggagcgggtaaccactcgtttct 2395
Db 2931 tccatggactaccgaagacagtgacttgggagtgaggagcgggtaaccactcgtttct 2990
QY 2396 ggtcactcctgagtgccagtcaccccttctaggtagagacttaactacaaagatgggagc 2455
Db 2991 ggtcactcctgagtgccagtcaccccttctaggtagagacttaactacaaagatgggagc 3050
QY 2456 tcaattttctttgaacaggaagacagagagagtgctgtgaataacaaacccactcagtg 2515
Db 3051 acaattttctttgaacaggaagacagagagagtgctgtgaataacaaacccactcagtg 3110
QY 2516 gttgacctccaaattagatgatgaatcgcactatacttcccgaagtaagcctgatca 2575
Db 3111 gttgacctccaaattagatgatgaatcgcactatacttcccgaagtaagcctgatca 3170
QY 2576 agatcacagtcctggttggagcaggtttcccgaagcctgggagaaacccagggatggg 2635
Db 3171 aaatacaattctggttggagcaggtttcccgaagcctgggagaaacccagggatggg 3230
QY 2636 tttggcaagcaagttcccacaggtatttcaactgaagcagtcagtcacacagtatc 2695
Db 3231 tttggcaagcaagttcccacaggtatttcaactgaagcagtcagtcacacagtatc 3290
QY 2696 agtcagacagtcaccccttgatagagaggtcgcagagaggaatttggccgcagtgctcaag 2755
Db 3291 agtcagacagtcaccccttgatagagaggtcgcagagaggaatttggccgcagtgctcaag 3350
QY 2756 attaatcaacagggcactcctagttcctgttccatcccttggataactcctcctgtacc 2815
Db 3351 attaatcaacagggcactcctagttcctgttccatcccttggataactcctcctgtacc 3410

QY 2816 ggttagaagacccctgggaccaatgattatcgaccagtacaggacttgagagaggtcaataa 2875
Db 3411 ggttagaagacccctgggactaatgactatcgaccagtacaggacttgagagaggtcaataa 3470
QY 2876 aaggggtgcaggaatacacacccacagctcccgaaaccccttataacccctttaggcgcctccc 2935
Db 3471 acgggtgcaggaatacacacccacagctcccgaaaccccttataacccctttaggcgcctccc 3530
QY 2936 gctcgaaagcaactggttacacagattgacttaaaagatgctcttctcctgcctgagatt 2995
Db 3531 acccaacagagctggtatcacagttatggacttaagagtgctcttctcctgcctgagatt 3590
QY 2996 acaccccaactagcccaacccacttttgccttcgaatggagatcccagggtacgggaagaac 3055
Db 3591 acaccccaactagcccaacccacttttgccttcgaatggagatcccagggtacgggaagaac 3650
QY 3056 cgggagctcaactggaccogactgcccgaaggttcaagaaccccagacacttttga 3115
Db 3651 cgggagctcaactggaccogactgcccgaaggttcaagaaccccagacacttttga 3710
QY 3116 cgaagccctacacagagacccctggccaaacttcaggaatcccaacccctcaggtgacccctc 3175
Db 3711 cgaagccctacacagagacccctggccaaacttcaggaatcccaacccctcaggtgacccctc 3770
QY 3176 ccagtaogtgatgaactgcttctggcgagcccaacaaacaggaactggttagaaggtac 3235
Db 3771 ccagtaogtgatgaactgcttctggcgagcccaacaaacaggaactggttagaaggtac 3830
QY 3236 gaaggaactactgctgaaattgctgacttagctacagagctctgctaaagagccca 3295
Db 3831 gaaggaactactgctgaaattgctgacttagctacagagctctgctaaagagccca 3890
QY 3296 gatttcagagagagagtaacttgggtacagtttcgcgggcgagcgatggct 3355
Db 3891 gatttcagagagagagtaacttgggtacagtttcgcgggcgagcgatggct 3950
QY 3356 gacggaggcagcgaagaaactgtagtccagataccggcccaacacacagcccaacaaagt 3415
Db 3951 gacggaggcagcgaagaaactgtagtccagataccggcccaacacacagcccaacaaagt 4010
QY 3416 gagagagttttgggagacagctgatttgcagacttgatcccgggttttgcgacctt 3475
Db 4011 gagagagttttgggagacagctgatttgcagacttgatcccgggttttgcgacctt 4070
QY 3476 agcagcccaactcaccctaaacaaagaaaggggagttctcctggcctcctgagca 3535
Db 4071 agcagcccaactcaccctaaacaaagaaaggggagttctcctggcctcctgagca 4130
QY 3536 ccagaaaggaatttgatgctatcaaaaagccctgtagcgacactgctctggccctccc 3595
Db 4131 ccagaaaggaatttgatgctatcaaaaagccctgtagcgacactgctctggccctccc 4190
QY 3596 tgacgttaacttaaacctttaccctttatggtgagcgttaagggagtagcagccgaggaat 3655
Db 4191 tgacgttaacttaaacctttaccctttatggtgagcgttaagggagtagcagccgaggaat 4250
QY 3656 ttttaaccccaacccctaggaccatggaggagacactgttgcctacctgtcaaaaagacttga 3715
Db 4251 ttttaaccccaacccctaggaccatggaggagacactgttgcctacctgtcaaaaagacttga 4310
QY 3716 tccgtgtagcagtggttggcccgatgctgaagcgtatcgagcgtgtggtccatactggt 3775
Db 4311 tccgtgtagcagtggttggcccgatgctgaagcgtatcgagcgtgtggtccatactggt 4370
QY 3776 caaggacgctgacaaattgactttgggacagataaactgtaagcccccaactgactt 3835
Db 4371 caaggacgctgacaaattgactttgggacagataaactgtaagcccccaactgactt 4430
QY 3836 ggagaaactcgttcgagccccccagacagatggatgacaaacgcccgcagatgacccacta 3895
Db 4431 ggagaaactcgttcgagccccccagacagatggatgacaaacgcccgcagatgacccacta 4490

CC histological localisation of activated retrovirus. Using Polymerase Chain
 CC Reaction DNA Quantitation (PDQ) on blood mononuclear cells. Infectivity
 CC titration and susceptibility testing can be performed. Ultimately animal
 CC donors without intact porcine retroviral sequences or with a lower copy
 CC number of viral elements could be selected.

XX Sequence 8060 BP; 2233 A; 1959 C; 2012 G; 1856 T; 0 other;

Query Match	60.7%;	Score 4986;	DB 18;	Length 8060;
Best Local Similarity	95.2%;	Pred. No. 0;		
Matches 5276;	Conservative 0;	Mismatches 245;	Indels 23;	Gaps 12;

QY	1	gtggtgtacgactgtggccccagcgcttggaataaaatctcttctgtgttgcac	60
DB	2530	gtggtgtacgactgtggccccagcgcttggaataaaatctcttctgtgttgcac	2589
QY	61	aagacgctctctcgtgagtgatt-tgggggtgcgtctctccagcccgagagggga	119
DB	2590	aagacgctctcgtgagtgattaaaggagtcgcttttccgagctggagg-----	2642
QY	120	ttgtctttacgtgctttcatttgggtgcgttggcgggaaatcctcgacacccctt	179
DB	2643	--ttcttttgcgtggtttacatttgggggctcgtcgggagtcgtcgccgccctta	2700
QY	180	acaccgagacacgacttgaggtga-aagggtaccctcttggacgtgtgtg--tgtgtc	236
DB	2701	acaccgagacacgacttgaggtgaataaaggatcctcttttaacgtgatgatgacc	2760
QY	237	ggcggcgctctgttctgagtgctgttttctggtgatgcgcttctcgtttgcagctg	296
DB	2761	ggcggcgctctgttctgagtgctgttttctggtgtgcgcttctcgtttgcagctg	2820
QY	297	tcctctcagaccgtgaagactgtaacagacagcgtgtaaggagatcacag	356
DB	2821	tcctctcagaccgtgaagactgtaacagacagcgtgtaaggagatcacag	2880
QY	357	gctgccacactggggagcggcgagggtgggagagcaggaagcgtgtgtgtctcc	416
DB	2881	gctgtgtcctggggagcggcgagggtgagagagcaggaagcgtgtgtgtctcc	2940
QY	417	tactgtcgtcagagacagcttctgtt-gttgaagcgaagcttccccctccggcg	476
DB	2941	tactgtcgtcagagacagcttctgtt-gttgaagcgaagcttccccctccggcg	3000
QY	477	tcgactcttttgctgt	536
DB	3001	tcgactcttttgctgt	3060
QY	537	gtttctgtctgt	596
DB	3061	gtttctgt	3120
QY	597	acagtgactaccccccttagttgactctcagaccattggactgaagtgatagtcagg	656
DB	3121	acagtgactaccccccttagttgactctcagaccattggactgaagtgatagtcagg	3180
QY	657	cataattgtcagttcaggttaagaaggacacttgacactttctgtctctgaaatgg	716
DB	3181	cataattgtcagttcaggttaagaaggacacttgacactttctgtctctgaaatgg	3240
QY	717	ccaacattcgttgt	776
DB	3241	cgacattcgttgt	3300
QY	777	gttaaggcaatcatttttcagactggacccgctctcactcctgtacagagccctatc	836
DB	3301	gttaaggcaatcatttttcagactggacccgctctcactcctgtacagagccctatc	3360
QY	837	cttacgtggcaagatttggcagaatctccgcctatgggttaaacatggcttaataa	896
DB	3361	cttacgtggcaagatttggcagaatctccgcctatgggttaaacatggcttaataa	3420

QY	897	ccaagaaacagggtcccggaatcctgtggtctgtgagagaaacaaacactcggcga	956
DB	3421	ccaagaaacagggtcccggaatcctgtggtctgtgagagaaacaaacactcggcga	3480
QY	957	aaagtgcagccctcttctcgttatcccccagagatcccgagcgcgacttggccga	1016
DB	3481	aaagtcaagccctc-tcctcatatcccccagagattgggaaccacgcttggccga	3539
QY	1017	acccaacactgttccccacaccccttatccagcacagaggtgctgtgagggga-cctctgc	1075
DB	3540	acccaacactgttccccacaccccttatccagcacagaggtgctgtgagggga-cctctgc	3599
QY	1076	cctctcgtgagctcgtgtgtgagggacacttctcgggactcggagccgagagagcgc	1135
DB	3600	cctctcgtgagctcgtgtgtgagggacacttctcgggactcggagccgagagagcgc	3659
QY	1136	caccgggagcgagacagcagatcgatataaccgctgcacactatggccctcccat	1195
DB	3660	caccgggagcgagacagcagatcgatataaccgctgcacactatggccctcccat	3719
QY	1196	gacgggggccaattgcagccctcagatttggcccttttctgtgagatctctataa	1255
DB	3720	acggggggccaattgcagccctcagatttggcccttttctgtgagatctctataa	3779
QY	1256	ttggaaaactaacacccctccttctcggagatcccaacgctcaggggttgggtga	1315
DB	3780	ttggaaaactaacaccccttctcggagatcccaacgctcaggggttgggtga	3839
QY	1316	gtcccttatgttctctcaccagcctacttgggatgtgtcaacagctgctgcagacact	1375
DB	3840	gtcccttatgttctctcaccagcctacttgggatgtgtcaacagctgctgcagacact	3899
QY	1376	cttcaaacaggagagcagagagaattctgttagagctagaaaaatgttctcgggc	1435
DB	3900	cttcaaacaggagagcagagagaattctgttagagctagaaaaatgttctcgggc	3959
QY	1436	cgacgggagccacagcagttgcataatgagattgacatggatttcccttgactcgc	1495
DB	3960	cgacgggagccacagcagttgcataatgagattgacatggatttcccttgactcgc	4019
QY	1496	cgttgggactacaacacgctgaaggttagggagcttgaataatctatcccgagctct	1555
DB	4020	cgttgggactacaacacgctgaaggttagggagcttgaataatctatcccgagctct	4079
QY	1556	ggtgaggggtccggggcgctcaagacggcccaactaatttggcctaaggttaagagagt	1615
DB	4080	ggtgaggggtccggggcgctcaagacggcccaactaatttggcctaaggttaagagagt	4139
QY	1616	gatcaggggacacgaacacccctcctcgttatttctgagagctcatggaagcctcag	1675
DB	4140	gatcaggggacacgaacacccctcctcgttatttctgagagctcatggaagcctcag	4199
QY	1676	goggttccccccttttgatctcctcagagggcccaagaaacccctcagtgggccctg	1735
DB	4200	goggttccccccttttgatctcctcagagggcccaagaaacccctcagtgggccctg	4259
QY	1736	catgtggcagctcggctctgatatcaggaagaacactcagagactgaaaggggttacagga	1795
DB	4260	catgtggcagctcggctctgatatcaggaagaacactcagagactgaaaggggttacagga	4319
QY	1796	ggctgagttacgtgatctagtgagagcagagagaaggtgtattacagaaggagacaga	1855
DB	4320	ggctgagttacgtgatctagtgagagcagagagaaggtgtattacagaaggagacaga	4379
QY	1856	agaggagaaggaacagagaaaaaggaaggaaggaaggaaggaaggaaggaaggaag	1915
DB	4380	agaggagaaggaacagagaaaaaggaaggaaggaaggaaggaaggaaggaaggaag	4439
QY	1916	acggcaagagagaagatttgaataagatttgcgcagctgtgttgaaggggaagcagcag	1975
DB	4440	acggcaagagagaagatttgaataagatttgcgcagctgtgttgaaggggaagcagcag	4499
QY	1976	ggagagagagagagatttttagggaaaattaggtcagggccctagacagtcagggaaac	2035

Db 4500 ggaagagagagagatttttaggaaaaataggtcaggccctagacagtcagggaacctggg 4559
Qy 2036 caataggaccactcgcagagaccagtgctgctattgttaagaaaaagacactgggc 2095
Db 4560 caataggaccactcgcagagaccagtgctgctattgttaagaaaaagacactgggc 4619
Qy 2096 aaggactgcccagaagaggaacaaagagaccgaaggtcctagctctagagaagataa 2155
Db 4620 aaggaaactgcccagaagaggaacaaagagaccgaaggtcctagctctagagaagataa 4679
Qy 2156 agattaggaggagacgggttcggaccctcccgagccaggggttaactttgaagggtgga 2215
Db 4680 agattaggaggagacgggttcggaccctcccgagccaggggttaactttgaagggtgga 4739
Qy 2216 ggggcaaacagttgacttccgttgatataccgagagcagatcattcagtgctgctacaacc 2275
Db 4740 ggggcaaacagttgacttccgttgatataccgagagcagatcattcagtgctactacagcc 4799
Qy 2276 attaggaactaaagaaaaaaatcctgggtgatgggtgccagagggcagcagta 2335
Db 4800 attaggaactaaagaaaaaaatcctgggtgatgggtg-cacagggcaacaacagta 4858
Qy 2336 tccatggactaccgaagacacttgacttggagtgaggagggtaacccactcgtttct 2395
Db 4859 tccatggactaccgaag-acagttgacttggagtgaggaggggttaacccactcgtttct 4917
Qy 2396 ggtcactcctgagtgcccagtcaccctctcaggtagagacttactgaccaagatgggagc 2455
Db 4918 ggtcactcctgagtgcccagtcaccctctcaggtagagacttattgaccaagatgggagc 4977
Qy 2456 tcaaatcttttgaaaggaagaccaggaagtgctgtgatacaacaacccactcactgt 2515
Db 4978 acaaatcttttgaaaggaagaccaggaagtgctgtgatacaacaacccactcactgt 5037
Qy 2516 gttgacctccaattagatgatgaatatcgactatatctcctcccaagtaaaagcctgatca 2575
Db 5038 gttgacctccaattagatgatgaatatcgactatatctcctccctagtaaaagcctgatca 5097
Qy 2576 agatatacagtgctgttgagcaggttccccaagcctgggcagaaaaaccgaggaagg 2635
Db 5098 aaatatacaattcgtgtggaacagttccccaagcctgggcagaaaaaccgaggaagg 5157
Qy 2636 ttggcaagaagttccccaagcctgttattcaactgaagccagtgctcacacagatc 2695
Db 5158 ttggcaagaagttccccaagcctgttattcaactgaagccagtgctcacacagatc 5217
Qy 2696 agtcagacagtcacccttgagtagagaggtcgcagaggaatttggccgcatgtccaag 2755
Db 5218 agtcagacagtcacccttgagtagagaggtcgcagaggaatttggccgcatgtccaag 5277
Qy 2756 attaatccaaagggaactcctagttcctgtccaatcccccttggaatactccccctgctacc 2815
Db 5278 attaatccaaagggaactcctagttcctgtccaatcccccttggaatactccccctgctacc 5337
Qy 2816 ggttaggaagcctgggacaaatgattatagacaggtacaggaacttgagagaggtccaataa 2875
Db 5338 ggttaggaagcctgggacaaatgattatagacaggtacaggaacttgagagaggtccaataa 5397
Qy 2876 aaggggtgagagacatacaccccaaggtcccgaaacctataaccttggagcgcctccc 2935
Db 5398 acgggtgagagatatacaccccaaggtcccgaaacctataaccttggagcgcctccc 5457
Qy 2936 gctgaacagggaactggtaacacagttattggaacttaaaagatgcttcttctcctgagatt 2995
Db 5458 accccaacaggagctggtatcacagttattggaacttaaaagatgcttcttctcctgagatt 5517
Qy 2996 aacccccactagccaacacacttttgccttcgaattggagagatccaggtacgggaagac 3055
Db 5518 aacccccactagccaacacacttttgccttcgaattggagagatccaggtacgggaagac 5577
Qy 3056 cgggagactcacctggaccctgctcccaaggttcaagaaactccccagaccatttga 3115

Db 5578 cgggagactcacctggaccgactgcccgaagggttcaagaaactccccagaccatttga 5637
Qy 3116 cgaagccctacacagagactggccaaactcagatcccaacccctcaggtgacccctct 3175
Db 5638 cgaagccctacacagagactggccaaactcagatcccaacccctcaggtgacccctct 5697
Qy 3176 ccagtcagtggtgactgctctcgtcgggagcccaaacacaggaactgcttagaaggtac 3235
Db 5698 ccagtcagtggtgactgctctcgtcgggagcccaaacacaggaactgcttagaaggtac 5757
Qy 3236 gaaggcaactactcgtggaattgtcactcaggtacagagcctctgcttaagaagccca 3295
Db 5758 gaaggcaactactcgtggaattgtcactcaggtacagagcctctgcttaagaagccca 5817
Qy 3296 gatttcagagagagagtaacatacttgggtgaggttgcggggcggcagcgatggct 3355
Db 5818 gatttcagagagagagtaacatacttgggtgaggttgcggggcggcagcgatggct 5877
Qy 3356 gacggaggcaggaagaaacttgatccagataccggcccccaaacacagcccaacaagt 3415
Db 5878 gacggaggcaggaagaaacttgatccagataccggcccccaaacacagcccaacaagt 5937
Qy 3416 gacagagtttttgagacagctgattttgcagactgtgatcccggttttgcacctt 3475
Db 5938 gacagagtttttgagacagctgattttgcagactgtgatcccggttttgcacctt 5997
Qy 3476 agcagccccactctacccgctaaacaaagaaaggggattctcctggctcctcagca 3535
Db 5998 agcagccccactctacccgctaaacaaagaaaggggattctcctggctcctcagca 6057
Qy 3536 ccaggaagcatttgatgctatacaaaaagccctgctgagcgaacctgctcctgcccctcc 3595
Db 6058 ccaggaagcatttgatgctatacaaaaagccctgctgagcgaacctgctcctgcccctcc 6117
Qy 3596 tgacgtaactaaaccccttaccctttatgttgatgagctaaaggagtagcccgagagt 3655
Db 6118 tgacgtaactaaaccccttaccctttatgttgatgagctaaaggagtagcccgagagt 6177
Qy 3656 tttaacccaaacctagagacattggaggagacctgttgcctcctgtcaaaagagcttga 3715
Db 6178 tttaacccaaacctagagacattggaggagacctgttgcctcctgtcaaaagagcttga 6237
Qy 3716 tccctgtagcagtggttgccctgctgctgaaggtatcgcagctgctgagcctactggt 3775
Db 6238 tccctgtagcagtggttgccctgctgctgagggctatcgagctgctgagcctactggt 6297
Qy 3776 caaggaagcgtgacaaattgactttggagc-agaataaactgtaataagcccccatgcat 3834
Db 6298 caaggaagcgtgacaaattgactttggagacaaataaactgtaataagcccccatgcat 6357
Qy 3835 tgagaaacactcgttcggcagccccccagacagatggatgacaaaccccgatgacccact 3894
Db 6358 tgagaaacactcgttcggcagccccccagacagatggatgacaaaccccgatgacccact 6417
Qy 3895 atcaaacctgctctcacagagaggtcacttgcctccacacagcgcctcctcaacctg 3954
Db 6418 atcaaacctgctctcacagagaggtcacttgcctccacacacagcgcctcctcaacctg 6477
Qy 3955 ccactctctgctgaagagactgataccagtgactcactgactcactgactcacttga 4014
Db 6478 ccactctctgctgaagagactgataccagtgactcactgactcacttga 6537
Qy 4015 ttgaggaactgggttcgcaagacacttacagacataccgctgactgaggaagtgttaa 4074
Db 6538 ttgaggaactgggttcgcaagacacttacagacataccgctgactgaggaagtgttaa 6597
Qy 4075 cctggttactgacggagcagctatgtgtgaaagtaagagtaggctggggcgagcag 4134
Db 6598 cctggttactgacggagcagctatgtgtgaaagtaagagtaggctggggcgagcag 6657
Qy 4135 tgggtgagcgggacccgacgatctggccagcagcctgcccgaaggaacttcagcgcaaa 4194
Db 6658 tgggtgagcgggacccgacgatctggccagcagcctgcccgaaggaacttcagcgcaaa 6717

QY 4195 aggtgagctcatggccctccagcaagcttttgcgctggtgaggaaggaatccataaaca 4254
 Db |||||
 QY 6718 aggtgagctcatggccctccagcaagcttttgcgctggtgaggaaggaatccataaaca 6777
 Db |||||
 QY 4255 ttatcagcagcagcaggtatgcttgcgactgcacagtgacacgggcccctctataaac 4314
 Db |||||
 QY 6778 ttatcagcagcagcaggtatgcttgcgactgcacagtgacatgggcccctataaac 6837
 Db |||||
 QY 4315 aaagggggtgctacctcagcagcaggggaaaaataaagaaacaaagagaaaattctaaacc 4374
 Db |||||
 QY 6838 aaagggggtgctacctcagcagcaggggaaaaataaagaaacaaagagaaaattctaaacc 6897
 Db |||||
 QY 4375 tattagaagccttacatttgcacaaaggttagctattatatactgtccttgacatcaga 4434
 Db |||||
 QY 6898 tattagaagcgtacatttaccacaaaggttagctattatatactgtccttgacatcaga 6957
 Db |||||
 QY 4435 aagcacaagatctctatatetagagggaaacccagatgctgacgggttgccaaagcagcag 4494
 Db |||||
 QY 6958 aagctaaagatctctatatccagaggaaacccagatgctgacgggttgccaaagcagcag 7017
 Db |||||
 QY 4495 ccagcgtgttaacctctgctataatagaaacgcgcacaaagcccaagcccaagacagac 4554
 Db |||||
 QY 7018 ccaggggttaacctctgctataatagaaatgcacaaagcccaagcccaagacagac 7077
 Db |||||
 QY 4555 agtacacctagaagactggcaagagataaaaaagatagaccagttctctgagactccgg 4614
 Db |||||
 QY 7078 agtacacctagaagactggcaagagataaaaaagatagaccatctctgagactccgg 7136
 Db |||||
 QY 4615 agggacacctgtatacctatattgggaaggaaatccctgccacacaaagaggggttagaat 4674
 Db |||||
 QY 7137 aaggacacctgtatacctcagattgggaaggaaatccctgccacacaaagaggggttagaat 7196
 Db |||||
 QY 4675 atgtccaac-agatacatcgtctaacccacactaggaactaaacacctgocagacttggtc 4733
 Db |||||
 QY 7197 atgtccacaagatatactgcttaacccacactaggaactaaacacctgocagacttggtc 7256
 Db |||||
 QY 4734 agaataccctctatcatgttctgagctaccagagtggtgctgactcgtggtgacaaact 4793
 Db |||||
 QY 7257 agaataccctctatcatgttctgagctaccagagtggtgctgactcgtggtgacaaact 7316
 Db |||||
 QY 4794 tgtgacctgcccagctggttaattgctaatctctccagataacctccagaaagagagacta 4853
 Db |||||
 QY 7317 tgtgacctgcccagctggttaattgctaatctctccagataacctccagaaagagagacta 7376
 Db |||||
 QY 4854 aggggaagccaccagcgtcactgggaagtgaactcactgaggttaaaagccggtctaaa 4913
 Db |||||
 QY 7377 aggggaagccaccagcgtcactgggaagtgaactcactgaggttaaaagccggtctaaa 7436
 Db |||||
 QY 4914 tacggaaacaaataatctattggtttttagacacaccttttcagagtggtgagagcttat 4973
 Db |||||
 QY 7437 tatggaaacaaataatctattggtttttagacacaccttttcagagtggtgagagcttat 7496
 Db |||||
 QY 4974 cctactaagaagagactcaacccgtggtgcttaagaaaaatactggaggaaatttttcca 5033
 Db |||||
 QY 7497 cctactaagaagagactcaacccgtggtgcttaagaaaaatactggaggaaatttttcca 7556
 Db |||||
 QY 5034 agattggaatacctaaagtaaggtcagacaaatggtccagctttcgttgcagagta 5093
 Db |||||
 QY 7557 agattggaatacctaaagtaaggtcagacaaatggtccagctttcgttgcagagta 7616
 Db |||||
 QY 5094 agtcaggagactggccaagatattgggattgattggaacctgcatgtgtacacagacc 5153
 Db |||||
 QY 7617 agtcaggagactggccaagatattgggattgattggaacctgcatgtgtacacagacc 7676
 Db |||||
 QY 5154 caaagctcaggacaggttagagagatgaatagaaacattaaagagacccttaccaaattg 5213
 Db |||||
 QY 7677 caaagctcaggacaggttagagagatgaatagaaacattaaagagacccttaccaaattg 7736
 Db |||||
 QY 5214 accacagagactggcctaataatgattgagctcctcctgcccctttgtgttttagggtg 5273
 Db |||||
 QY 7737 acccggagagactggcctaataatgattgagctcctcctgcccctttgtgttttagggtg 7796
 Db |||||

QY 5274 aggaacacccctggacagtttgggtgactgaaccccttatgaattgctctacgggggacccccc 5333
 Db |||||
 QY 7797 aggaacacccctggacagtttgggtgactgaaccccttatgaattgctctacgggggacccccc 7856
 Db |||||
 QY 5334 ccgttggcagaaattgctttgcacatagtgctgagtgctgctttcccaagcctttgttc 5393
 Db |||||
 QY 7857 ccattgtagaaattgcttctgtacatagtgctgagtgctgctttcccaagcctttgttc 7916
 Db |||||
 QY 5394 tctaggtcgaagcgtcgtgagtggtgagcagcgtggaagcagctcccgaggacc 5453
 Db |||||
 QY 7917 tctaggtcgaagcgttgggtgagcagcagcgtggaagcagctcccgaggacc 7976
 Db |||||
 QY 5454 tactc---agaggaagacttgcacatgctccacatgcttccaaagttagagattcagcttat 5510
 Db |||||
 QY 7977 tactcagagagagagacttgcacatgctccacatgcttccaaagttagagattcagcttat 8036
 Db |||||

RESULT 12

AAAF77725

ID AAF77725 standard; cDNA; 8060 BP.

XX AC AAF77725;

XX AC AAF77725;

DT 23-MAY-2001 (first entry)

XX Tsukuba-1 cDNA.

XX Tsukuba-1; retrovirus; graft transplantation; xenotransplantation; ss.

XX Porcine retrovirus.

XX US6190861-B1.

XX 20-FEB-2001.

XX 13-DEC-1996; 96US-0766528.

XX 14-DEC-1995; 95US-0572645.

XX (GEHO) GEN HOSPITAL CORP.

XX Fishman JA;

XX WPI; 2001-256211/26.

Assessing risk of endogenous retroviruses in clinical practice and in
 xenotransplantation, comprises using probe sequences derived from swine
 or miniature swine retroviral genome -
 Claim 1; Fig 1; 127pp; English.
 The present invention relates to a method for screening a cell or tissue
 for the presence or expression of a retrovirus (RV), comprising
 contacting a target nucleic acid from the cell or tissue with a second
 nucleic acid from the present invention (e.g. the present sequence or a
 fragment thereof). The method is useful for RV detection and to assess
 graft transplantation risk. Screening of animals allows the elimination
 of donors with active replication of known viruses. Inactive proviruses
 can be detected and inactivated, allowing identification and elimination
 of potential human pathogens derived from swine in a manner not possible
 in the outbred human organ donor population and is important to the
 development of human xenotransplantation.

Sequence 8060 BP; 2233 A; 1959 C; 2012 G; 1856 T; 0 other;
 Query Match 60.7%; Score 4986; DB 22; Length 8060;
 Best Local Similarity 95.2%; Pred. No. 0;
 Matches 5276; Conservative 0; Mismatches 245; Indels 23; Gaps 12;

|||||
Db 4680 agattagggagacgggttcgagcccccctcccgagcccgaggttaactttgaagtgga 4739
QY 2216 ggggcaaccagttagttcctggttgataccggagcggagcattcaagtgcgtacaacc 2275
Db 4740 ggggcaaccagttagttcctggttgataccggagcggagcattcaagtgcgtacaacc 4799
QY 2276 attaggaactaaaagaaaaaactcctgggtgatgggtgcacagggcaacggcagta 2335
Db 4800 attaggaactaaaagaaaaaactcctgggtgatgggtgcacagggcaacggcagta 4858
QY 2336 tccatggactaccgaagaacccgttgactggagtgggagcgggtaaaccaactcgtttct 2395
Db 4859 tccatggactaccgaag - acagttgacttggagtgggacgggtaaaccaactcgtttct 4917
QY 2396 ggtcatccctgagtgccagtagtaccctcttagttagagacttactgaccaagatggagc 2455
Db 4918 ggtcatccctgagtgccagtagtaccctcttagttagagacttactgaccaagatggagc 4977
QY 2456 tcaaatcttttgaaacaaggaagaccagaagtgtctgtgaataacaaacccatcactgt 2515
Db 4978 acaaatcttttgaaacaaggaagaccagaagtgtctgtgaataacaaacccatcactgt 5037
QY 2516 gttgacctccaattagatgaatatcagactatatctccccaagttaagcctgatca 2575
Db 5038 gttgacctccaattagatgaatatcagactatatctccccaagttaagcctgatca 5097
QY 2576 agatatacagtcctggttgagcagtttcccaagcctggcgagaacccgagggatggg 2635
Db 5098 aatatacaattctgttggaacagtttcccaagcctggcgagaacccgagggatggg 5157
QY 2636 ttgtgcaagcaagttcccccacaggttatccaaactgaagccagtgctacacacagtatc 2695
Db 5158 ttgtgcaagcaagttcccccacaggttatccaaactgaagccagtgctacacacagtatc 5217
QY 2696 agtcagacgtaccccttgtagtagagagcctcgagaggaatttggcccatgttcaag 2755
Db 5218 agtcagacgtaccccttgtagtagagagcctcgagaggaatttggcccatgttcaag 5277
QY 2756 attaatccaaagggatccctagtcctctgccaatcccttggaataactccctgtacc 2815
Db 5278 attaatccaaagggatccctagtcctctgccaatcccttggaataactccctgtacc 5337
QY 2816 ggttagaagcctgggaccacttatcgaccagtagcacgacttagagaggtcaataa 2875
Db 5338 ggttagaagcctgggactaatgactatcgaccagtagcacgacttagagaggtcaataa 5397
QY 2876 aaggttgacagacatacacccacggtcccgaaacccctataacccctttgagcgccctcc 2935
Db 5398 acgggtgcaggaatacacccacagtcctcggaacccctataacccctttgtgtctccc 5457
QY 2936 gcctgaacggaaactggttacacagatttgacttaaaagatgcctcttctgctgagatt 2995
Db 5458 acccaacggagctggtatacagatttgacttaaaagatgcctcttctgctgagatt 5517
QY 2996 acaccccaactagccaacacttttgcctcgaaatggagatccagggtacgggaagaac 3055
Db 5518 acaccccaactagccaacacttttgcctcgaaatggagatccagggtacgggaagaac 5577
QY 3056 cgaggagctcaactgagccagactgcccaagggttcaagaactccccagacatctttga 3115
Db 5578 cgaggagctcaactgagccagactgcccaagggttcaagaactccccagacatctttga 5637
QY 3116 cgaagccctacacaggaactggccaaacttcaggatcccaacccctcaggtgacctcct 3175
Db 5638 cgaagccctacacaggaactggccaaacttcaggatcccaacccctcaggtgacctcct 5697
QY 3176 ccagtagctggatgactcgtcttgccgggagcccaacacaggaactgcttagaaggtac 3235
Db 5698 ccagtagctggatgactcgtcttgccgggagcccaacacaggaactgcttagaaggtac 5757
QY 3236 gaaggcaactactcgtggaatttctgacgtaggtacagagcctctgtataaaggccca 3295
|||||

Db 5758 gaaggcaactactcgtggaatttctgacactaggtacagagcctctgtataaagaagccca 5817
QY 3296 gatttcagggagagaggttaacatacttgggttacagtttgcggggcgagcagatggct 3355
Db 5818 gatttcagggagagaggttaacatacttgggttacagtttgcggggcgagcagatggct 5877
QY 3356 gacggaggcacgaagaaaaactgtagtcagataccgcccacacacagcccaaacagt 3415
Db 5878 gacggaggcacgaagaaaaactgtagtcagataccgcccacacacagcccaaacagt 5937
QY 3416 gagagagtttttggggacagctggattttgcagactgtgatcccggtgttttgcacctt 3475
Db 5938 gagagagtttttggggacagctggattttgcagactgtgatcccggtgttttgcacctt 5997
QY 3476 agcagcccaactctacccgtacccaaagaaaaaggggattctcctgggtccttgagca 3535
Db 5998 agcagcccaactctacccgtacccaaagaaaaaggggattctcctgggtccttgagca 6057
QY 3536 ccagaagggcatittgatgtctatacaaaagccctgctgagcagcactgctctggccctcc 3595
Db 6058 ccagaagggcatittgatgtctatacaaaagccctgctgagcagcactgctctggccctcc 6117
QY 3596 tgacgttaactaaacccctttacccctttatgtggtgagcgttaagggttagcccgaggagt 3655
Db 6118 tgacgttaactaaacccctttacccctttatgtggtgagcgttaagggttagcccgaggagt 6177
QY 3656 ttttaacccaaaccccttaggacattgaggagacactgttgccctacctgtcaaaagaaacttga 3715
Db 6178 ttttaacccaaaccccttaggacattgaggagacactgttgccctacctgtcaaaagaaacttga 6237
QY 3716 tctgtagccagtggttgcccgctatgtctgaagcctatcgagcagctgtgcccatactggt 3775
Db 6238 tctgtagccagtggttgcccgctatgtctgaagcctatcgagcagctgtgcccatactggt 6297
QY 3776 caagcagctgacaaatgaacttttgggac - agaataactgtataatagccccccatgcat 3834
Db 6298 caagcagctgacaaatgaacttttgggacaaagataaactgtataatagccccccatgcat 6357
QY 3835 tgagaaacatcgttcggcagccccagaccagatgtagacaaacgcccgcagatgacccact 3894
Db 6358 tgagaaacatcgttcggcagccccagaccagatgtagacaaacgcccgcagatgacccact 6417
QY 3895 atcaagcctgcttctcacagaggggtcaacttctgcctccaccagcctctcaacccctg 3954
Db 6418 atcaagcctgcttctcacagaggggtcaagctgcgtccaccacagcctctcaacccctg 6477
QY 3955 ccactcttctgctgaagagactgatgaaccagtgactcatgattgccatcaacttga 4014
Db 6478 ccactcttctgctgaagagactgatgaaccagtgactcatgattgccatcaacttga 6537
QY 4015 ttgaggagactggggtccgcaagacccctacagacatcccgactgactgagagatgctaa 4074
Db 6538 ttgaggagactggggtccgcaagacccctacagacatcccgactgactgagagatgctaa 6597
QY 4075 cctggttcaactgacggaagcagctatgtggtggaaggttaaagaggtggtggtggcgag 4134
Db 6598 cctggttcaactgacggaagcagctatgtggtggaaggttaaagaggtggtggtggcgag 6657
QY 4135 tggtagagcgggacccgacgactcctgggacgagcctcccggaaggaactcaagcga 4194
Db 6658 tggtagagcgggacccgacgactcctgggacgagcctcccggaaggaactcaagcga 6717
QY 4195 aggcgtgactcatggcctccagcaagcttgcggctgagcgaagggaaatccataaca 4254
Db 6718 aggcgtgactcatggcctccagcaagcttgcggctgagcgaagggaaatccataaca 6777
QY 4255 tttatcacggacagcaggtatgcctttgcagctgcacacgtacacgggcccactctataaac 4314
Db 6778 tttatcacggacagcaggtatgcctttgcagctgcacacgtacacgggcccactctataaac 6837
QY 4315 aaaggggttgccttacctcagcagggagggaaataaagaaacaaagaggaatcttaagcc 4374
Db 6838 aaaggggttgccttacctcagcagggagggaaataaagaaacaaagaggaatcttaagcc 6897

Qy 4375 tattagaagccttacatttgcacaaaggctagctattatataactgtcctggacacaga 4434
Dy 6898 tattagaagcgttacatttaccacaaaggctagctattatataactgtcctggacacaga 6957
Qy 4435 aagcacaagatctcatatagagggaacacagatgctgacccgggttgccaaagcagcag 4494
Dy 6958 aagctaaagatctcatatccagaaacacagatgctgacccgggttgccaaagcagcag 7017
Qy 4495 cccaggctgttaacctctcctataataaagcccaagcccaagcccaagcccaagcccaagcagc 4554
Dy 7018 cccagggtgttaacctctcctataataaagcccaagcccaagcccaagcccaagcagcagc 7077
Qy 4555 agtacacctagaagactggcaagagataaaagatagaccaggttctctgagactccgg 4614
Dy 7078 agtacacctagaagactggcaagagataaaagatagacca-ttctctgagactccgg 7136
Qy 4615 aggggacctgtatcatctatctatgggaagaaatcctgcccccaaaagaggggttagaat 4674
Dy 7137 aagggacctgtatcatctcagatgggaagaaatcctgcccccaaaagaggggttagaat 7196
Qy 4675 atgtccaac-agatacatctcctacccacactagaaactaaacacctgcagcagttggtc 4733
Dy 7197 atgtccacaagatcacatctcctacccacactagaaactaaacacctgcagcagttggtc 7256
Qy 4734 agaacaatcccttatcatgttctgaggtaccagagatggtgactcgtggtgcaaacat 4793
Dy 7257 agaacaatcccttatcatgttctgaggtaccagagatggtgactcgtggtgcaaacat 7316
Qy 4794 tgtgtccctgcagcagtggttaatgctaaactctccagaatacctccaggaagagacta 4853
Dy 7317 tgtgtccctgcagcagtggttaatgctaaactctccagaatgctccaggaagagacta 7376
Qy 4854 aggggaagccacccagcgcctcactgggaagtgactcactcagagtgaaagccggtctaaa 4913
Dy 7377 aggggaagccacccagcgcctcactgggaagtgactcactcagagtgaaagccggtctaaa 7436
Qy 4914 taccgaaacaaatctattgtttttgttagacacccctttccaggtggttagaggtctat 4973
Dy 7437 tatggaacaaatctattgtttttgttagacacccctttccaggtggttagaggtctat 7496
Qy 4974 cctactaagaagagacttcaacgctggtgctaaagaaataactggaggaatttttcca 5033
Dy 7497 cctactaagaagagacttcaacgctggtgctaaagaaataactggaggaatttttcca 7556
Qy 5034 agatttgaatacctaaagtgataggttcagacaaatggtccagctttcgttgcacaggtta 5093
Dy 7557 agatttgaatacctaaagtgataggttcagacaaatggtccagcttttgttgcacaggtta 7616
Qy 5094 agtcagggactggccaagatattggggatgattggaactgcaattgtgcatacagacc 5153
Dy 7617 agtcagggactggccaagatattggggatgattggaactgcaattgtgcatacagacc 7676
Qy 5154 caaagctcagacaggtagagagatgaatagaaccattaaagaccccttaccacaaattg 5213
Dy 7677 caaagctcagacaggtagagagatgaatagaaccattaaagaccccttaccacaaattg 7736
Qy 5214 accacagagactggcaataatgattggatggctcctcctcctttgtgcttttttaggggtg 5273
Dy 7737 accgagagactggcgttaattgattggtatgctcctcctcctttgtgcttttttaggggtt 7796
Qy 5274 aggaacacccctggacagtttggctgacccctctatgaattgctctacagggggacccccc 5333
Dy 7797 aggaacacccctggacagtttggctgacccctctatgaattgctctacagggggacccccc 7856
Qy 5334 ccggttgcagaaattgccttgcacatagctgctgctgcttgccttccagccctttgttc 5393
Dy 7857 ccattggttagaaattgcttctgtacatagctgctgctgcttccagccctttgttc 7916
Qy 5394 tctaggtcctcaggtcagttggtgagcagcagcagcgtggaagcagcgtccggagggcc 5453
Dy 7917 tctaggtcctcaggtcagttggtgagcagcagcagcgtggaagcagcgtccggagggcc 7976

Qy 5454 tactc---aggaggagacttgcacatccttcccaagttggagattcagcttat 5510
Dy 7977 tactcaggaggaggagacttgcagatccacatcgtttcccaagttggagattcagcttat 8036
Qy 5511 gttagacgcccgcgtgcaggaac 5534
Dy 8037 gttagacgcccgcgtgcaggaac 8060
RESULT 13
AAC67020
ID AAC67020 standard; DNA; 4402 BP.
XX
AC AAC67020;
XX
DT 27-MAR-2001 (first entry)
XX
DE PERV env protein coding sequence SEQ ID NO: 20.
XX
KW Xenotransplantation; infectious agent; vaccine; ds.
XX
OS Porcine endogenous retrovirus.
XX
PN WO200071726-A1.
XX
PD 30-NOV-2000.
XX
PF 24-MAY-2000; 2000WO-US14296.
XX
PR 24-MAY-1999; 99US-0135631.
XX
PA (MAYO-) MAYO MEDICAL VENTURES.
XX
PI Federspiel MJ;
XX
DR WPI; 2001-032041/04.
XX
PT Inhibiting or preventing infectious agent transmission in mammalian
transplant recipients, by introducing recombinant DNA comprising DNA
encoding extracellular proteins of the agent into donor cells, such as
swine cells -
XX
PS Claim 16; Page 105-106; 144pp; English.
XX
CC The present invention provides a method to prevent the transmission of
infectious agents during xenotransplantation. This involves introducing
to donor swine cells a recombinant DNA encoding a peptide fragment from
the infectious agent, and then introducing these cells into the
transplant recipient.
XX
SQ Sequence 4402 BP; 1259 A; 1085 C; 1111 G; 947 T; 0 other;

Query Match 45.7%; Score 3755.4; DB 22; Length 4402;
Best Local Similarity 92.3%; Pred. No. 0;
Matches 4042; Conservative 0; Mismatches 271; Indels 64; Gaps 6;

Qy 1940 gatcttggccgcagtggttgaaggagagcagcagcagggagagagagatttagaa 1999
Dy 1 gatcttggctgcagtggttgaagggaacaaacacacagagagagatttagaa 60
Qy 2000 aattaggtcagccctagacagtcagggaacctgggcaatagaccctcagacaaga 2059
Dy 61 aattaggtcagccctagacagtcagggaacctgggcaatagaccctcagacaaga 120
Qy 2060 ccaggtgcgtattttaaagaaagacactgggcaaggaactgcccaagaggaa 2119
Dy 121 ccaatgtgcattattttaaagaaagacactgggcaaggaactgcccaagaggaa 180
Qy 2120 caaaggaccacaggtccttagcttagaagaagataagattagggagacgggttcga 2179
Dy 181 caaaggactgaaggtcttagcttagaagaagataagattagggagacgggttcga 240

XX Key Location/Qualifiers
 PH 911..2884
 FT CDS
 FT /*tag= a
 FT /product= envelope_protein

XX WO9853104-A2.

XX 26-NOV-1998.

XX 18-MAY-1998; 98WO-GB01428.

XX 16-MAY-1997; 97GB-0010154.

XX (MEDI-) MEDICAL RES COUNCIL.

XX Stoye JP, Weiss RA;

XX WPI; 1999-045324/04.

XX P-PSDB; AAW85452.

XX Newly isolated nucleic acid probe capable of hybridising to either the PERV-A or PERV-B env gene - useful in the detection of retroviruses, and their subtypes, in a sample of porcine/human tissue

XX Claim 3; Page 21-23; 36pp; English.

CC The present sequence encodes a Pig endogenous retrovirus (PERV)-B envelope protein. PERV exists in two different subtypes, PERV-A and PERV-B. The differences are reflected in sequence divergence in the envelope genes. Probes and primers can be derived from the envelope (env) genes of PERV-A and PERV-B. The probes and primers are used in a method to detect retroviruses in a sample of porcine/human tissue, particularly primary porcine tissue and human cell lines that have been cultivated in the presence of a porcine cell line, or human tissue from a patient with a xenotransplant. Subtype of PERV in a sample containing one of the PERV env genes can also be determined.

XX Sequence 3482 BP; 927 A; 854 C; 867 G; 834 T; 0 other;

Query Match 41.1%; Score 3376.6; DB 20; Length 3482;
 Best Local Similarity 99.0%; Pred. No. 0;
 Matches 3441; Conservative 0; Mismatches 29; Indels 7; Gaps 4;

QY 4718 cctgcagcagttgtgcagaacatcccttatcatgttctgagctaccagagtggtcga 4777
 |||||
 Db 6 cctgcagcagttgtgcagaacatcccttatcatgttctgagctaccagagtggtcga 65
 QY 4778 ctgggtgttcaaacattgtgtgcctgcagctggttaattgttaattcttccagaatacc 4837
 |||||
 Db 66 ctgggtgttcaaacattgtgtgcctgcagctggttaattgttaattcttccagaatacc 125
 QY 4838 tccagagaaagagactaagggaaggaacaccagcgctcactggaagtggaacttactga 4897
 |||||
 Db 126 tccagagaaagagactaagggaaggaacaccagcgctcactggaagtggaacttactga 185
 QY 4898 ggttaagccggctaaatcacggaacaaataatctattgtttttgtagacaccttttcag 4957
 |||||
 Db 186 ggttaagccggctaaatcacggaacaaataatctattgtttttgtagacaccttttcag 245
 QY 4958 atgggtgaggttactactactaagaaagagacttcaacogtgggtggctaaagaaatact 5017
 |||||
 Db 246 atgggtgaggttactactactaagaaagagacttcaacogtgggtggctaaagaaatact 305
 QY 5018 ggaagaaatttttccagattttggaatacctaaggttaattggtcagacaatggtccagc 5077
 |||||
 Db 306 ggaagaaatttttccagattttggaatacctaaggttaattggtcagacaatggtccagc 365
 QY 5078 ttctgttgcaggaagtagcagagactggccagagattgtgggattgttggaactgca 5137
 |||||
 Db 366 ttctgttgcaggaagtagcagagactggccagagattgtgggattgttggaactgca 425

QY 5138 ttgtgcatacagaccccaagctcagagcaggtagagaggtgaatagaacattaaaga 5197
 |||||
 Db 426 ttgtgcatacagaccccaagctcagagcaggtagagaggtgaatagaacattaaaga 485
 QY 5198 gaccttaccaaattgaccacagagactggcattaatgattggatggctctctcgtccctt 5257
 |||||
 Db 486 gaccttaccaaattgaccacagagactggcattaatgattggatggctctctcgtccctt 545
 QY 5258 tgtgttttttaagggtgaggaacacccctggacagttttgagctgagccccccttaagtgtct 5317
 |||||
 Db 546 tgtgttttttaagggtgaggaacacccctggacagttttgagctgagccccccttaagtgtct 605
 QY 5318 ctacgggggacccccccgttggcagaaattgcctttgcacatagctgctgctgtct 5377
 |||||
 Db 606 ctacgggggacccccccgttggtagaaattgcttctgtacatagctgctgctgtct 665
 QY 5378 ttcccagcctttgttctctaggtcgaagcgtcagtggtggagcgagcagcgtgaa 5437
 |||||
 Db 666 ttcccagcctttgttctctaggtcgaagcgtcagtggtggagcgagcgtgaa 725
 QY 5438 gacgtccgggagggcctactactcagg--aggagacttgaagtccaactgcttccaagt 5494
 |||||
 Db 726 gacgtccgggagggcctactcaggagaagagacttgaagtccaactgcttccaagt 785
 QY 5495 tggagattcagctctatgttagacgccacgtgcaggaacacctcgagactcgtgtgaaagg 5554
 |||||
 Db 786 gggagattcagctctatgttagacgccacgtgcaggaacacctcgagactcgtgtgaaagg 845
 QY 5555 acctatctcgtacttttgaccacacacacgctgtgaaagtgcaggaatccccacctg 5614
 |||||
 Db 846 ccttattctcgtacttttgaccacacacacgctgtgaaagtgcaggaatccccacctg 905
 QY 5615 gatcatgcacccccacgttaagccggccacacctcccgactcgggtgagagcgacgaaaa 5674
 |||||
 Db 906 gatcatgcacccccacgttaagccggccacacctcccgactcgggtgagagcgacgaaaa 965
 QY 5675 gactgagaatcccccttaagcttcgctccatcgctggttcttactactacaataactc 5734
 |||||
 Db 966 gactgagaatcccccttaagcttcgctccatcgctggttcttactactacaataactc 1025
 QY 5735 cccagggcagtagtaaacgccttatagacagctcgaaccccccatagacatttatccctta 5794
 |||||
 Db 1026 cccagggcagtagtaaacgccttatagacagctcgaaccccccatagacatttatccctta 1085
 QY 5795 cctgctgattattgaacctgatacgggtgtcaactgtaataagactcagaggtgtgtctc 5854
 |||||
 Db 1086 cctgctgattattgaacctgatacgggtgtcaactgtaataagactcagaggtgtgtctc 1145
 QY 5855 ctgagggcacctcgttggcctgaactgttctgctccgattgtattaccacccgtgtta 5914
 |||||
 Db 1146 ctgagggcacctcgttggcctgaactgttctgctccgattgtattaccacccgtgtta 1205
 QY 5915 aaagcacacacctcccaaccttagtccgttagttgggtttctattgtgccaggacagaga 5974
 |||||
 Db 1206 aaagcacacacctcccaaccttagtccgttagttgggtttctattgtgccaggacagaga 1265
 QY 5975 aagagaatactgttgggggttctggggaatcctctgttaggagatggagctgctcaact 6034
 |||||
 Db 1266 aagagaatactgttgggggttctggggaatcctctgttaggagatggagctgctcaact 1325
 QY 6035 ccaacgagtaggagactggaatggcgaatctctccagaccgggtgaaatctctcttgg 6094
 |||||
 Db 1326 ccaacgagtaggagactggaatggcgaatctctccagaccgggtgaaatctctcttgg 1385
 QY 6095 tcaattccggccggcggaagtagcaaaatgatgaaactatataaagataagagctgctccc 6154
 |||||
 Db 1386 tcaattccggccggcggaagtagcaaaatgatgaaactatataaagataagagctgctccc 1445
 QY 6155 catcagacttagattatctaaagataagtttcaactg--aaagaaaaacaggaaaaatttc 6212
 |||||
 Db 1446 catcagacttagattatctaaagataagtttcaactgaaagaaagaaaaacaggaaaaatttc 1505

QY 6213 aaaagtggataaaatggtatgactgctggggaatagttttttat-tatatggggggagcag 6271
DB 1506 aaaagtggataaaatggtatgactgctggggaatagtttttttaaaatagttggggggagcag 1565
QY 6272 ggtccactttaaccattcgcttagatagacgaggggacagaaacccctgtggcaatgg 6331
DB 1566 ggtccactttaaccattcgcttagatagacgaggggacagaaacccctgtggcaatgg 1625
QY 6332 gacccgataaaagtactggtgaaacagggccccggccccctggagccaccgcataaactgc 6391
DB 1626 gacccgataaaagtactggtgaaacagggccccggccccctggagccaccgcataaactgc 1685
QY 6392 cgggtcccccaattaaacctcgctgccccgcataaaacacagcccccttagcaaacagtacca 6451
DB 1686 cgggtcccccaattaaacctcgctgccccgcataaaacacagcccccttagcaaacagtacca 1745
QY 6452 ctgattgattctccacaaacgctgaaactccccaggtgttcctgttaagacagagac 6511
DB 1746 ctgattgattctccacaaacgctgaaactccccaggtgttcctgttaagacagagac 1805
QY 6512 agagactcttcagttctatccagggagctttcccaagccatacaactccaccgacccctgatg 6571
DB 1806 agagactcttcagttctatccagggagctttcccaagccatacaactccaccgacccctgatg 1865
QY 6572 ccactttcttgggttctgtctatctcctcagggcctccttattatgaggggatgcta 6631
DB 1866 ccactttcttgggttctgtctatctcctcagggcctccttattatgaggggatgcta 1925
QY 6632 aagaaagaaaattcaattgtgacaaagagacatgaaatacaatgtacatgggggtcccgaa 6691
DB 1926 aagaaagaaaattcaattgtgacaaagagacatgaaatacaatgtacatgggggtcccgaa 1985
QY 6692 ataagcttacccctcactgaaagtctccgggaaggagacatgcataggaagagctcccccac 6751
DB 1986 ataagcttacccctcactgaaagtctccgggaaggagacatgcataggaagagctcccccac 2045
QY 6752 cccaccacaccttctgtatagttctgtttagagcagggcctcagaaatacagttatt 6811
DB 2046 cccaccacaccttctgtatagttctgtttagagcagggcctcagaaatacagttatt 2105
QY 6812 tagtacctgggtataacaggtgggtgggcatgcaatacctgggttaacccctgtgttccaa 6871
DB 2106 tagtacctgggtataacaggtgggtgggcatgcaatacctgggttaacccctgtgttccaa 2165
QY 6872 cctcagttctcaacacatccaaagatttctgtctatggttccaaatcgtcccccgagtg 6931
DB 2166 cctcagttctcaacacatccaaagatttctgtctatggttccaaatcgtcccccgagtg 2225
QY 6932 actaccatcctgaggaagtgttctctgataaataagagagctgttagcaacccctagaag 6991
DB 2226 actaccatcctgaggaagtgttctctgataaataagagagctgttagcaacccctagaag 2285
QY 6992 gagaacccgtatcccttaccctagctgtaattgctcgattaggaagccggttggcgtag 7051
DB 2286 gagaacccgtatcccttaccctagctgtaattgctcgattaggaagccggttggcgtag 2345
QY 7052 gaacagggagagctgccttgatcacaggaacacagcagctagagaaggacttggtagc 7111
DB 2346 gaacagggagagctgccttgatcacaggaacacagcagctagagaaggacttggtagc 2405
QY 7112 tacatcgccgcatgacagaagatctccagagcccttaaaagagctgttagcaacccctagaag 7171
DB 2406 tacatcgccgcatgacagaagatctccagagcccttaaaagagctgttagcaacccctagaag 2465
QY 7172 agtccctgacttcttctgtcgaagtgttctacagaacccggaggttagatctcgtgt 7231
DB 2466 agtccctgacttcttctgtcgaagtgttctacagaacccggaggttagatctcgtgt 2525
QY 7232 ttctaagagaggttgggttattgtcagccttaaaagaagaattgttcttctatgtagatc 7291
DB 2526 ttctaagagaggttgggttattgtcagccttaaaagaagaattgttcttctatgtagatc 2585
QY 7292 actcaggagccatcagagactccatgaaacagcttagaagaaagtttagagagcgctcgaa 7351

DB 2586 actcaggagccatcagagactccatgagcaagcttagaagaggttagagagcgctcgaa 2645
QY 7352 ggaagagagaggtgaccaggggtgttgaagatggttcaacaggtctccttgatga 7411
DB 2646 ggaagagagaggtgaccaggggtgttgaagatggttcaacaggtctccttgatga 2705
QY 7412 ccaccctgcttctgctcagcgggccccctagtagtctcctgttacttaacagtg 7471
DB 2706 ccaccctgcttctgctcagcgggacccctagtagtctcctgttacttaacagtg 2765
QY 7472 ggccttgcttaataataggtttgttgcctttgttagagaacgagtgagtcacag 7531
DB 2766 ggccttgcttaataataggtttgttgcctttgttagagaacgagtgagtcacag 2825
QY 7532 tcatgatacttagcacaacagtagtcccaagggccttctgagcaagagagaactgacctagc 7591
DB 2826 tcatgatacttagcacaacagtagtcccaagggccttctgagcaagagagaactgacctagc 2885
QY 7592 ctccccagttctaaagattagaactattaacaagacaagaagtggggaatgaagagatgaa 7651
DB 2886 ctccccagttctaaagattagaactattaacaagacaagaagtggggaatgaagagatgaa 2945
QY 7652 aatgcaacctaaacctccacagaccaggaagttaataaaaaagcttaaatgcccccga 7711
DB 2946 aatgcaacctaaacctccacagaccaggaagttaataaaaaagcttaaatgcccccga 3005
QY 7712 ttacagacctgctggtcctcagtagtaaataggtaggaaggtcacacttctattgttccag 7771
DB 3006 ttacagacctgctggtcctcagtagtaaataggtaggaaggtcacacttctattgttccag 3065
QY 7772 gctgctatcctggcctaaagataacacagaaatgagttgactaatacgttatcttggaa 7831
DB 3066 gctgctatcctggcctaaagataacacagaaatgagttgactaatacgttatcttggaa 3125
QY 7832 ttctgtaaaactgactggcacacatagaagaattgattacacattgacagccctagtgacc 7891
DB 3126 ttctgtaaaactgactggcacacatagaagaattgattacacattgacagccctagtgacc 3185
QY 7892 tatctcaactgcaatctgctcactctccacagagcccccagatgcggacccctcgagc 7951
DB 3186 tatctcaactgcaatctgctcactctccacagagcccccagatgcggacccctcgagc 3245
QY 7952 tattttaaaatgattggttccacggagcgcggtcctcgatatttttaaaatgattggtcca 8011
DB 3246 tattttaaaatgattggttccacggagcgcggtcctcgatatttttaaaatgattggtcca 3305
QY 8012 tggagcgcggtc-tcgatatttttaaaatgattggttggacgacagccttctgtgt 8070
DB 3306 cggagcgcggtcctcctcgatatttttaaaatgattggttggacgacagccttctgtgt 3365
QY 8071 gaacccataaaagctgtcccgattccgactcggcgggcgccgagtccttacccttgcgtg 8130
DB 3366 gaacccataaaagctgtcccgattccgactcggcgggcgccgagtccttacccttgcgtg 3425
QY 8131 gtgtacgactgtggccccagcgcttgggaataaaaaatcctctgtgttgcac 8187
DB 3426 gtgtacgactgtggccccagcgcttgggaataaaaaatcctctgtgttgcac 3482

RESULT 15

AAC67022

ID AAC67022 standard; DNA; 4918 BP.

XX AAC67022;

XX AAC67022;

DT 27-MAR-2001 (first entry)

DE PERV env protein coding sequence SEQ ID NO: 22.

XX Xenotransplantation; infectious agent; vaccine; ds.

XX Porcine endogenous retrovirus.

Search completed: February 24, 2002, 03:02:54
 Job time: 19340 sec

```

Db      3789  agaaagggttagatcgttattcttaaaagaagggttatgtgtagccttgaaggag 3848
OY      7270  gaatgtgcttatactagatcctcagagccatcagagctcactgaacaagcttaga 7329
Db      3849  gaatgctgtttatgtgtatcattcagagggccatcagagactccatgaacaagcttaga 3908
OY      7330  aaaaatltagaagagcgtcgaaggaagaagagcgtcgaaggggtgtgttgaagatgtg 7389
Db      3909  gaagaggttgagagagcgtcgaagggaaagaaactactcaaggggtgtgttgaagagatgg 3968
OY      7390  ttcaaaaggtctccttgaatgaaccacccgtcttctgtctcgaaggggccccctagatgc 7449
Db      3969  ttcaaaaggtctccttgaatgaaccacccgtcttctgtctcgaaggggccccctagatgc 4028
OY      7450  ctgtcctgttactcaagtttggtggttgaatgaatgaatgaatgaatgaatgaatgaatga 7509
Db      4029  ctccctcgttactcaagtttggtggttgaatgaatgaatgaatgaatgaatgaatgaatga 4088
OY      7510  gaacgaatgagtcagtcacaga tcatgttacttagcacaagctacaaagccctcctagac 7569
Db      4089  gaacgaatgagtcagtcacaga tcatgttacttagcacaagctacaaagccctcctagac 4148
OY      7570  caaggaagaactgaaccttagccttccagttcctaagattagaactatlaacaagaacag 7629
Db      4149  ---agagaagcgtgcgcgttagcttaccagttcctaagattagaactatlaacaagaagaag 4205
OY      7630  aagtgaggatgaagaatgaagaatgaacccaacccaacccaacccaacccaacccaaccca 7689
Db      4206  aagtgaggatgaagaatgaagaatgaacccaacccaacccaacccaacccaacccaaccca 4265
OY      7690  aaaaagctcctaaatgcccccgaaatlaacagaccctgtcgtgcagtaaatagatgaagg 7749
Db      4266  agaaagctcctaaatgcccccgaaatlaacagaccctgtcgtgcagtaaatagatgaagg 4317
OY      7750  tcaacacttccattgttccagggcctgtatccctgtgcctgaagtaagaacaagaatga 7809
Db      4318  tcaacacttctgtgtgttgaagggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 4377
OY      7810  gttaga-----ctaatc 7820
Db      4378  aaaaacagggccctgtgatatgtcctctatgtcttgaacactcttgaacactgtccttaact 4437
OY      7821  gcttatcgtgattctgttaaacctgactg----- 7848
Db      4438  gctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 7849
OY      7849  --gcaccatagaagaatgtatgaacatltgaacagccctagtgacctatctcaactgaa 7905
Db      4498  aacggaaccccgatgaatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 4557
OY      7906  tctgtactctgcacgaagcccaacgacatgcgagaccctcggagcctatltlaaatgat 7965
Db      4558  catgtctcccccaccccgaaacatgcaaatgtgttaactctaaacaattlaaatlaaat 4617
OY      7966  tggfccaagggagcggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 8025
Db      4618  tggfccaagggagcggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 4664
OY      8026  tcatatltlaaaatgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 8085
Db      4665  --gatatttgaatgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 4722
OY      8086  tgtccgattccgactcgcgggcgcgacgtccttaacccctgcgtgtgtgtgtgtgtgtgtgt 8145
Db      4723  tgtccgactcgcacactcgcgggcgcgacgtccttaacccctgcgtgtgtgtgtgtgtgtgt 4782
OY      8146  ccccaagcgcgcttggaataaaatccctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 8191
Db      4783  ccccaagcgcgcttggaataaaatccctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 4828
  
```



```

FT      protein_bind      7885..7914
FT      /*tag= p
FT      /note= "Binding site for transcription factors
FT      AP-1/GATA, c-Myc and AP-1"
FT      protein_bind      7928..7953
FT      /*tag= q
FT      /note= "Binding site for transcription factors E47
FT      and ETS-1"
FT      CAAT_signal      7964..7968
FT      /*tag= r
FT      /note= "Binding site for transcription factor GATA"
FT      protein_bind      7985..7997
FT      /*tag= s
FT      /note= "Binding site for transcription factor GATA"
FT      CAAT_signal      8003..8007
FT      /*tag= t
FT      protein_bind      8024..8036
FT      /*tag= u
FT      /note= "Binding site for transcription factor GATA"
FT      CAAT_signal      8042..8046
FT      /*tag= v
FT      protein_bind      8049..8060
FT      /*tag= w
FT      /note= "Binding site for transcription factor
FT      AP-1/CREB"
FT      TATA_signal      8075..8090
FT      /*tag= x
FT      misc_feature      1..61
FT      /*tag= y
FT      /note= "R LTR domain"
FT      misc_feature      8127..8135
FT      /*tag= z
FT      /note= "Polya downstream element"
FT      polyA_signal      8161..8166
FT      /*tag= aa
FT      W09740167-A1.
FT      30-OCT-1997.
FT      PD
FT      18-APR-1997; 97WC-GB01087.
FT      PF
FT      10-FEB-1997; 97GB-0002668.
FT      PR
FT      19-APR-1996; 96GB-0008164.
FT      XX
FT      PA
FT      (IMUT-) IMUTRAN LTD.
FT      (QONE-) Q-ONE BIOTECH LTD.
FT      XX
FT      Galbraith DN, Haworth C, Lees GM, Smith KT;
FT      WI: 1997-535851/49.
FT      DR
FT      P-PSDB; AAM39271; AAM39272; AAM39273.
FT      XX
FT      Polynucleotide encoding porcine retrovirus expression product -
FT      useful to develop products for use in vaccines, diagnosis and
FT      xeno-transplantation
FT      PT
FT      Claim 4; Fig 3; 69pp; English.
FT      PS
FT      XX
FT      This DNA sequence encodes the porcine retrovirus (PoRV) virion core
FT      CC polypeptide (GAG), polymerase (POL), and envelope (ENV) proteins and
FT      CC also includes the Long Terminal Repeat (LTR). These proteins can be used
FT      CC to develop viral vaccines, antisense nucleic acids, ribozymes and other
FT      CC antiviral agents. They can also be used in xeno-transplantation
FT      CC technology and as diagnostic tools.
FT      XX
FT      Sequence 8209 BP; 2168 A; 2064 C; 2154 G; 1823 T; 0 other;

```

alignment_scores:
 Quality: 2801.00 Length: 524
 Ratio: 5.345 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-171-553B-4 x AAV09700 ..
 Align seg 1/1 to: AAV09700 from: 1 to: 8209

```

1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisTrpH 17
|||||
588 ATGGGACAGACAGTGTACCTACCCCTTAGTTGACCTCGACCATTTGAC 637
17 rGluValArgSerArgAlaHisAsnLeuSerValGlnValLysGlyP 34
|||||
638 TGAAGTTAGATCCAGGCTCATATTTGTCAGTTGACGTTTACAGAGGAC 687
34 rOTrpgInThrPheCysAlaSerGluTrpProThrPheAspValGlyTrp 50
|||||
688 CTGGGACAGACTTCTGTGCTCTGAATGGCCACATTCGATGTTGATGAG 737
51 ProSerGluGlyThrPheAsnSerGluLleLeuAlaValLysAla11 67
|||||
738 CCATCAGAGGGGACCTTAAATTCGAATATTCCTGCTGTTAAGGCAAT 787
67 eLlePheGlnThrGlyProGlySerHisProAspGlnGluProTrpLle 84
|||||
788 CATTTTCAGACTGAGCCGCTCTCATCTGATCAGAGCCCTATATTC 837
84 eutHrTrpGlnAspLeuAlaGluAspProProProTrpValLysProTrp 100
|||||
838 TTACGTGCGCAAGATTGGCAGAGATCCGCCCATGGGTTAAACCATG 887
101 LeuAsnLysProArgLysProGlyProAlaGlyLleLeuAlaLeuGlyLys 117
|||||
888 CTAATTAACCAAGAAAGCCAGGTCCTCCGATCTGCTTGGAGAGAA 937
117 sAsnLysHisSerAlaGluLysValGluProSerSerSerTrpLeuPro 134
|||||
938 AAACAAACACTCGGCCGCAAAAAGTCAGCCCTTCTCTGCTATCTACCC 987
134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150
|||||
988 GAGATCGAGAGCGCGCGACTTGCGCGAACCACCACTTCCGCCACCC 1037
988 GAGATCGAGAGCGCGCGACTTGCGCGAACCACCACTTCCGCCACCC 1037
151 ProLeuSerSerThrGlyCysGlyGlyThrSerAlaProProGlyAla 167
|||||
1038 CCTTATCCAGACAGGCTGCTGAGGGGACCTGCCCCCTCTGAGAG 1087
167 aProValValGluGlyProAlaAlaGlyThrArgSerArgGlyAla17 184
|||||
1088 TCCGGTGTGGAGGAGACTCTGCGCGAGACTCGGAGCCGAGAGGCGCCA 1137
184 hrProGluArgThrAspGluLleAlaLleLeuProLeuArgThrTyGly 200
|||||
1138 CCCGAGAGCGGACAGACAGATGCGGATATTACCGCTGCGCACCTATGCG 1187
1138 CCCGAGAGCGGACAGACAGATGCGGATATTACCGCTGCGCACCTATGCG 1187
201 ProPrometProGlyGlnLeuGlnProLeuGlnTyTrpProPheSe 217
|||||
1188 CTTCCATGCGCAGGGGCGCAATTGCGACCCCTCAGATTTGGCCCTTTTC 1237
1188 CTTCCATGCGCAGGGGCGCAATTGCGACCCCTCAGATTTGGCCCTTTTC 1237
217 rSerAlaAspLeuTyAsnTrpLysThrAsnHisProProPheSerGln 234
|||||
1238 TTTCGCGAGATCTCATATATTGGAATACTACCATCCCTTCTGCGAGG 1287
1238 TTTCGCGAGATCTCATATATTGGAATACTACCATCCCTTCTGCGAGG 1287
234 sPProGlnArgLeuThrGlyLeuValGluSerLeuMetPheSerHisGln 250
|||||
1288 ATCCCAAGCGCTCAGCGGGGTGGTGAGTCCCTTATGTTCTCTCAGCAG 1337
1288 ATCCCAAGCGCTCAGCGGGGTGGTGAGTCCCTTATGTTCTCTCAGCAG 1337
251 ProThrTrpAspAspCysGlnGlnLeuLeuGlnThrLeuPheThrArg 267
|||||
1338 CCTACTTGGGATGATTTGTCACAGCTGCTGACACACTTTCACACCGCA 1387
1338 CCTACTTGGGATGATTTGTCACAGCTGCTGACACACTTTCACACCGCA 1387
267 uGluArgGluArgLleLeuLeuGluAlaArgLysAsnValProGlyAla 284
|||||
1388 GGAACGAGAGAGAAATTTCTTTAGAGGCTAGAAAAAATGTTCTCTGGGCGC 1437

```

```

284 spgIyArpProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLeu 300
1438 ACGGGCAGCCAGCTGCAAAATGATTCATGGAGTTCCCTTG 1487
301 ThrArpGlyTrpAspPyrAsnThrAlaGluGlyArgGluSerLeu 317
1488 ACTGGCCCCGGGTGGAGTACACACGCTGAAAGTAGAGGAGAGCTTGA 1537
317 sILTrpArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgRpr 334
1538 AATCATGCCAGGCTGTGGTGGGCTCCGGGGCCCTCAAGACGCG 1587
334 rorhrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGlnPro 350
1588 CCACTAATTTGGCTAAGTAGAGAGGATGATCGAGGCGCAAGAACT 1637
351 ProSerValPheLeuGluArgLeuMetGluAlaPheArgArgPheThr 367
1638 CCTCGGTATTTCTTGAGAGGCTCATGAAAGCTTCAGGGGCTTCA 1687
367 oPhasProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 384
1688 TTTGATCTCTAGCTCAGAGGCCCGCAAGAGCTCAGTGGCTGGCTTCA 1737
384 IeGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGly 400
1738 TTGGGCACATGGCTCTGATATCAGAAAGAACTTCAGAGACTGGAAGG 1787
401 LeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGlyVal 417
1788 TTACAGGAGGCTGAGTTACGTGATCTAGTGAGAGAGGCGAGAGTGT 1837
417 rTyArgArgGluThrGluGluGluGluGluGluGluGluGluGlu 434
1838 TTACAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1887
434 rGgGluArgGluGluArgArgAspArgGlnGluGluGluGluGlu 450
1888 GAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1937
451 LysIleLeuAlaAlaValAlaGluGlyLysSerSerArgGluArg 467
1938 AAGATCTTGGCCGAGTGTGAAGGAGAGAGAGAGAGAGAGAGAG 1987
467 gAspPheArgLysIleArgSerGlyProArgGlnSerGlyAsnLeu 484
1988 AGATTTTGGAAAATTAGGTAGGCTCCTAGACAGCTCAGGGAGCT 2037
484 snArGThrProLeuAspLysAspGlnCysAlaIyrCysLysGlu 500
2038 ATAGAGACCCAGCTCGACAGAGAGAGAGAGAGAGAGAGAGAGAG 2087
501 HisTrpAlaArgAsnGlyProLysLysGlyAsnLysGlyProLys 517
2088 CACTGGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2137
517 uAlaLeuGluGluAspLysAsp 524
2138 AGCTCTAGAGAGAGATTAAGAT 2159
seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:AAV09699
seq_documentation_block:
ID AAV09699 standard; DNA; 8196 BP.
XX
AC AAV09699;
XX
DT 19-MAY-1998 (first entry)
XX
DE Porcine retrovirus DNA encoding, GAG, POL and ENV.
XX
KW Porcine retrovirus; POEV; POL protein; ENV protein; GAG protein;
vaccine; diagnosis; xenotransplantation; prophylactic; therapeutic; ds.

```

```

XX Porcine retrovirus.
OS
XX
FH Key Location/Qualifiers
FT CDS 576..2126
FT CDS /tag=a
FT /product=GAG protein
FT /note="vifion core polypeptide"
FT CDS 2143..5733
FT /tag=b
FT /product=POL protein
FT /note="polymerase peptide sequence as given in
specification"
FT CDS 506..7576
FT /tag=c
FT /product=ENV protein
FT /note="envelope protein"
XX
PD WO9740167-A1.
XX
PF 30-OCT-1997.
XX
PF 18-APR-1997; 97WO-GB01087.
XX
PR 10-FEB-1997; 97GB-0002668.
PR 19-APR-1996; 96GB-0008164.
XX
PA (IMOT-) IMOTRAN LTD.
PA (QONE-) Q-ONE BIOTECH LTD.
XX
PI Galbraith DN, Haworth C, Lees CM, Smith KT;
XX
DR WPI: 1997-535851/49.
XX
PT Polynucleotide encoding porcine retrovirus expression product -
PT useful to develop products for use in vaccines, diagnosis and
PT xeno-transplantation
XX
PS Claim 4; Fig 2; 6pp; English.
XX
CC This DNA sequence encodes the porcine retrovirus (POEV) vifion core
CC polypeptide (GAG), polymerase (POL) and envelope (ENV) proteins. These
CC proteins can be used to develop viral vaccines, antisense nucleic acids,
CC ribozymes and other antiviral agents. They can also be used in
CC xeno-transplantation technology and as diagnostic tools.
XX
SQ Sequence 8196 BP; 2165 A; 2061 C; 2147 G; 1820 T; 3 other;

alignment_scores:
Quality: 2774.00 Length: 524
Ratio: 5.304 Gaps: 0
Percent Similarity: 99.809 Percent Identity: 99.809

alignment_block:
US-09-171-553B-4 x AAV09699 ..
Align seg 1/1 to: AAV09699 from: 1 to: 8196

1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisTrpH 17
|||||
576 ATGGAGACAGACAGTACCTACCCCTTACTTGGACTCTGACCATTTGGAC 625
|||||
17 rGluValArgSerArgAlaHisAsnLeuSerValGlnValLysLysGlyP 34
|||||
626 TGAAGTTAGATCCAGGCTCATTAATTGTTCAGTTCAAGTTAAGAGGAG 675
|||||
34 rorrrglnhrPheCysAlaSerGluTrpProThrPheAspValGlyTrp 50
|||||
676 CTGGGACAGACTTCTGCTCTGAAATGGCCACATTCATGTTGATGG 725
|||||
51 ProSerGluGlyThrPheAsnSerGluIleIleLeuAlaValLysAlaI 67
|||||

```

```

726 CCATCAGAGGGGACCTTAAATTCGAATTAATCTGGCTGTTAAGCAAT 775
67 e11eeheclnhrGlyProGlySerHisProaspGlnuProGlyIlel 84
|||||
776 CATTTTTCAGACTGGACCCGCTCATCTGATCAGAGAGCCCTATATCC 825
84 euThrTrpGlnaspLeuAlaGluaspProProProTrpValLysProTrp 100
|||||
826 TTACCTGGCAGATTGGCAGAGATCCCTCCGCACTGGGTTAAACCAAGG 875
101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaGlnLysGly 117
|||||
876 CTAATTAACCAAGAGAACCCAGTCCCGAATCCGCTCTTGGAGAGAA 925
117 SasLysHisSerAlaGlnLysValGluProSerSerSerTyrLeuPro 134
|||||
926 AAACAAACACTCGCCGAAAAAGTCGAGCCCTCTCTCTGATCTACCCC 975
134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150
|||||
976 GAGATCGAGAGAGCCGCGACTTGGCGGAGACCCCAACCTGTTCCCCACC 1025
151 ProLeuSerSerThrGlyCysCysGluGlyThrSerAlaProProGlyAl 167
|||||
1026 CCTTATTCAGACAGGGGTGCTGTGAGGGAGACCTGCGCTCTCGAGC 1075
167 aProValValGluGlyProAlaAlaGlyThrArgSerArgArgLysAla 184
|||||
1076 TCCGCTGTGTGAGGAGCTCTGCGCGGAGTCGAGCGGAGAGGCGCA 1125
184 hrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrGly 200
|||||
1126 CCCCAGACGAGACAGAGATCGGATTAACCGCTGCGCACTATAGGC 1175
201 ProPrometProGlyGlnLysGlnLeuGlnProLeuGlnTyrTrpProPhe 217
|||||
1176 CCTCCAGAGCAGGGGGCCATTCAGAGCCCTCCAGATATGGCCCTTTTC 1225
217 rSerAlaAspLeuTyrAsnTrpLysThrAsnHisProProPheSerGlu 234
|||||
1226 TTCTGCAGATCTCTATTAATGTGAAACATAACCATCCCTTTCTCGAGG 1275
234 sPProGlnArgLeuThrGlyLeuValGluSerLeuMetPheSerHisGln 250
|||||
1276 ATCCCAAGCGCTCAGCGGGGTGTGAGTCCCTTATGTCTCTCCACAG 1325
251 ProThrTrpAspAspCysGlnGlnLeuLeuGlnThrLeuPheThrThrG 267
|||||
1326 CCTACTTGGGATGATGTCAACACCTGCTGCAGACACTCTTCACAAACGA 1375
267 uGluArgGluArgIleLeuLeuGlnAlaArgLysAsnValProGlyAla 284
|||||
1376 GGACCGAGAGAGAAATCTGTAGAGGCTAGAAAAAATGCTCGGGGCGG 1425
284 sPGLyArgProThrGlnLeuGlnAsnGluLeuAspMetGlyPheProLeu 300
|||||
1426 ACGGGCCAGCCACCTGCAAAATGATGATGACATGGAGATTCCTCTTG 1475
301 ThrArgProGlyLysProTyrAsnThrAlaGlnLysArgLysLeuLys 317
|||||
1476 ACTGCGCCCGGTGGACTACAAACAGCGTGAAGGTAGGAGAGCTTGAA 1525
317 sIleTyrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArg 334
|||||
1526 AATCTATCGCAGGCTCTGTGGGGGTCTCGGGGGCGCTCAAGAGGCG 1575
334 roThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGluPro 350
|||||
1576 CCACTAATTTGGCTAAGTAAGAGAGGTGATGCAAGGAGAGCAAGCACT 1625
351 ProSerValPheLeuGluArgLeuMetGluAlaPheArgArgPheThrP 367
|||||
1626 CCTCTGGATTTCTTGTGAGAGGCTCATGAAAGCCCTTCAGGGGGTTCAACCC 1675

```

```

367 oPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPheI 384
|||||
1676 TTTTGATCTTAACCTCAGAGGCCCAAGAAAGCCCTAGTGGCCCTGCTTCA 1725
384 leGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGlnGly 400
|||||
1726 TTGGGAGTGGCTGTGATATCAGGAAGAACTTCAGAGACTGTGAAGG 1775
401 LeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGlnLysValTy 417
|||||
1776 TTACAGAGAGCTGAGTTACTGATCTTGTGACAGAGCGAGAGAGTGTA 1825
417 rTyrArgArgGluThrGlnGlnLysGlnLysGlnArgLysGlnLysGlu 434
|||||
1826 TTACAGAGAGGAGACAGAAAGAGAGAGAGAACAGAAAAAAGAGAGAG 1875
434 rGGLuGluArgGluGluArgArgAspArgArgGlnGlnLysAsnLeuThr 450
|||||
1876 GAGAGAGAAAGGAGAGAAAGACGTGATAGACGGCAAGAGAAATTTGACT 1925
451 LysIleLeuAlaAlaValAlaGlnLysSerSerArgGluArgGluArg 467
|||||
1926 AAGATCTTGGCCCGCAGTGTGTAAGGAGAGAGACGACGAGGAGAGAG 1975
467 gaSPheArgLysIleArgSerGlyProArgGlnSerGlyAsnLeuGly 484
|||||
1976 AGATTTAGAGAAATTAAGTCAGGCGCTGACAGCTCAGGAGAACTGGCA 2025
484 sNArgThrProLeuAspLysAspGlnCysAlaIleTyrCysLysGlnLysGly 500
|||||
2026 ATAGAGCCCTCAGCTGACAAAGAGACAGTGTGCGTATGTGAAGAAAAAG 2075
501 HisTrpAlaArgAsnCysProLysLysGlyAsnLysGlyProLysVal 517
|||||
2076 CACTGGGCAAGAGACTGCCCAAGAGAGAGAAACAAAGAGACCAAGTTC 2124
517 uAlaLeuGlnGluAspLysAsp 524
|||||
2125 ACCTCTAGAAAGAGATAAGAT 2146

seq_name: /SID2/gcdata/geneseq/geneseq/NR2001.DAT:AAF77726
seq_documentation_block:
ID   AAF77726 standard; DNA; 7333 BP.
XX
AC   AAF77726;
XX
DT   23-MAY-2001 (first entry)
XX
DE   Defective retroviral genome isolated from PK-15 cell line.
XX
KW   Retrovirus; graft transplantation; xenotransplantation; PK-15 cell line;
XX
SS   .
XX
OS   unidentified.
XX
PN   US6190861-B1.
XX
PD   20-FEB-2001.
XX
PP   13-DEC-1996; 96US-0766528.
XX
PR   14-DEC-1995; 95US-0572645.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
XX
PI   Fishman JA;
XX
WI   WI: 2001-256211/26.
XX
DR   P-PSDB; AAB73282, AAB73283, AAB73284.
XX
PT   Assessing risk of endogenous retroviruses in clinical practice and in

```

PT xenotransplantation, comprises using probe sequences derived from swine
 PT or miniature swine retroviral genome -
 XX
 PS Claim 1; Fig 2; 127pp; English.
 XX

CC The present invention relates to a method for screening a cell or tissue
 CC for the presence or expression of a retrovirus (RV), comprising
 CC contacting a target nucleic acid from the cell or tissue with a second
 CC nucleic acid from the present invention (e.g. the present sequence or a
 CC fragment thereof). The method is useful for RV detection and to assess
 CC graft transplantation risk. Screening of animals allows the elimination
 CC of donors with active replication of known viruses. Inactive proviruses
 CC can be detected and inactivated, allowing identification and elimination
 CC of potential human pathogens derived from swine in a manner not possible
 CC in the outbred human organ donor population and is important to the
 CC development of human xenotransplantation.

XX Sequence 7333 BP; 1984 A; 1788 C; 1885 G; 1676 T; 0 other;

alignment_scores:

Quality: 2758.00 Length: 525
 Ratio: 5.263 Gaps: 1
 Percent Similarity: 99.810 Percent Identity: 99.619

alignment_block:

US-09-171-553B-4 x AAF77726 ..

Align seg 1/1 to: AAF77726 from: 1 to: 7333

1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisThrPth 17
 598 ATGGGACAGACAGTACTACCCCTTAGTTGACTTCGACCATTTGAGC 647
 17 rGluValArgSerArgAlaHisAsnLeuSerValGlnValLysLysGly 34
 648 TGAAGTTAGATCCAGGGCTCATTAATTTGTCACTTCAAGGTTAAGAGGAG 697
 34 rOTrGlnThrPheCysAlaSerGluTrpProThrPheAspValGlyTrp 50
 698 CTTGGCAGACTTTCGTGCTCTGAATGGCCACATTCGATTTGGATGG 747
 51 ProSerGluGlyThrPheAsnSerLutIleIleuAlaValLysAlaI 67
 748 CCATCAGAGGGGACCTTAATCTGAATTAATCTCGCTGCTTAAGCAAT 797
 67 eIlePheGlnThrGlyProGlySerHisProAspGlnGluProTyrIle 84
 798 CATTTTCAGACTGAGCCGGCTCTCATCTGATCAGAGGCCCTATATCC 847
 84 euThrTrpGlnAspLeuAlaGluAspProProProTrpValLysProTr 100
 848 TTACGTGGCAAGATTGGCAGAGATCTCCGCCATGGGTTAAACCATGG 897
 101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGlu 117
 898 CTAATAATAACCAAGAAAGCCAGAGTCCCGAATCTGCTTGGAGAGAA 947
 117 sAsnLysHisSerAlaGluLysValGluProSerSerSerLysLeuPro 134
 948 AAACAACACCTCGCCGAAAGAGTGAACCCCTC.TCCCTGATCTACACCC 996
 134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150
 997 GAGATCGAGAGACCCCGCAGCTTGGCCGGAACCCCACTGTTCCCCACC 1046
 151 ProLeuSerSerThrGlyCysCysGluGlyThr.SerAlaProProGly 167
 1047 CCTTATCCAGACAGGGTGTGTAAGGGGACCTCTGCCCTCCTGAG 1096
 167 laProValValGluGlyProAlaAlaGlyThrArgSerArgGlyAla 183
 1097 CTCGGTGTGAGGAGACCTCTGCCGGGACTCGAGCCGGAAGAGCCGCC 1146

184 ThrProGluArgPheAspGluIleAlaIleLeuProLeuArgThrTyrG 200
 1147 ACCCGGAGCGGACAGACAGATCGCATATTACCGCTGCCACTATGG 1196
 200 yProPoweMetProGlyGlnLeuGlnProLeuGlnTyrTrpProPhe 217
 1197 CCTTCCCATTCGACAGGGGCCAATTGACGCCCTCCAGATATGGCCCTTT 1246
 217 eSerAlaAspLeuTyrAsnTrpLysThrAsnHisProProPheSerGlu 233
 1247 CTTTCGAGATCTCTATTAATGAAACTAATCAATCCCTTTCTCCGAG 1296
 234 AspProGlnArgLeuThrGlyLeuValGlnSerLeuMetPheSerHisG 250
 1297 GATCCCAACCCCTCAGCGGGTGTGTGAGTCCCTTATGTTCTTCACCA 1346
 250 nProThrTrpAspAspCysGlnGlnLeuLeuGlnThrLeuPheThrTh 267
 1347 GCCTACTTGGGATGATTGTCAACAGCTGTCTGACAGACTCTTCACACCG 1396
 267 LuGluArgGluArgIleLeuLeuGlnAlaArgLysAsnValProGlyAla 283
 1397 AGGAGCGAGAGAGAAATCTTTAGAGGCTAAAAAAATGTTCCCGGGCC 1446
 284 AspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProle 300
 1447 GACGGCGGACCCACGCACTTCCAAATGAGATTACATTCGATTTCCCTT 1496
 300 uThrArgProGlyTrpAspTyrAsnThrAlaGluGlyArgGluSerLeu 317
 1497 GACTGCGCCCGGTTGGGACTTACACAGCGCTGAGAGGTAGGAGAGCTTGA 1546
 317 ySIIeTyrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArg 333
 1547 AAATCTATCGCCAGGCTCTGGTGGGGGTCTCCGGGGCGGCTCAAGACGG 1596
 334 ProThrAsnLeuAlaLysValArgIleValMetGlnLysProAsnGluPr 350
 1597 CCCCAATATTGGCTTAAGGTAGAGAGGTATGAGAGGACCGAAGCAAC 1646
 350 oProSerValPheLeuGlnArgLeuMetGluAlaPheArgArgPheThr 367
 1647 TCCCTCGATTTCTTGTAGAGAGCTCATGAAAGCCCTCAGCGGTTACCC 1696
 367 ropPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 383
 1697 CTTTGTATCTTACCTCAGAGGCCCGAAGACCTCAGTGGCCCTGGCTTC 1746
 384 IleGlyInserAlaLeuAspIleArgLysLysLeuGlnArgLeuGlu 400
 1747 ATTGGGAGTCCGCTCTGGATATTCAGAGAAACCTTGAAGACTGGAGG 1796
 400 yLeuGlnGluAlaGluLeuValArgAspLeuValArgGluAlaGlyVal 417
 1797 GTTACAGAGAGGTGATGATTACGTGATTAAGAGAGGACGAGAGAGTGT 1846
 417 yTyrTArgArgGluThrGluGluGluLysGluGlnArgGlyGluLysGlu 433
 1847 ATTACAGAAAGGAGACAGAAAGAGAGAAACAGAGAAAAAGAGAG 1896
 434 ArgGluGluArgGluGluArgArgAspArgArgGluGluLysAsnLeuTh 450
 1897 AGAGAAAGAAAGGAGAGAAAGCTGATAGACGGCAGAGAAAGATTGGAC 1946
 450 rLysIleLeuAlaAlaValValGluGlyLysSerSerArgGluArgGlu 467
 1947 TAAAGTCTTGCCGCACTGTTGAAGGAGAGACAGAGGAGAGAGAGAG 1996
 467 rGAspPheArgLysIleArgSerGlyProArgGlnSerGlyAsnLeuGly 483
 1997 GAGATTTTAGAAATATAGTTCAGGCCCTTAGACAGTACGAGGAACCTGGC 2046

```

484 AsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGlyLysGly 500
|||||
2047 AATGAGACCCACCTCGACACAGACAGCTGCTGATTTGTAAGAAAAAG 2096
500 YHSTRPAlaArgAsnGlyProLysGlyGlyAsnLysGlyProLysVal 517
|||||
2097 ACACGGGCGCAAGCACTGCCCAAGAAAGGCAACAAAGCAAGCAAGCTCC 2146
517 euAlaLeuGluGluAspLysAsp 524
|||||
2147 TAGCTCTAGAAAGATTAAGAT 2169

seq_name: /SID2/9/cgdata/geneseq/geneseq/NA1997.DAT.AAT74883
seq_documentation_block:
ID AAT74883 standard; cDNA; 7393 BP.
XX
AC AAT74883;
XX
DT 09-FEB-1998 (first entry)
XX
DE Porcine retrovirus cDNA (defective).
XX
KW Retrovirus; porcine; GAG protein; POL protein; ENV protein;
KW xenotransplantation; infectious; provirus; organ transplant; donor;
KW activated virus; PCR; ss.
XX
OS Porcine retrovirus.
XX
FH Key Location/Qualifiers
FT CDS 598..2172
FT /tag= a
FT mat_peptide 598..2169
FT /tag= b
FT CDS 2320..4737
FT /note= "putative GAG protein"
FT /tag= c
FT /note= "putative POL coding region (partial) as
FT /note= "described in the specification"
FT mat_peptide 2320..3522
FT /tag= d
FT /note= "putative POL protein (partial)"
FT 3516..4328
FT /tag= e
FT /note= "putative POL protein (partial)"
FT mat_peptide 4332..4748
FT /tag= f
FT CDS 4738..6725
FT /tag= g
FT /note= "putative ENV coding region (partial) as
FT /note= "described in the specification"
FT mat_peptide 4752..6722
FT /tag= h
FT /note= "ENV protein (partial)"
XX
PN WO9721836-A1.
XX
PD 19-JUN-1997.
XX
PF 13-DEC-1996; 96MO-US19680.
XX
PR 14-DEC-1995; 95US-0572645.
XX
PA (GENO ) GEN HOSPITAL CORP.
XX
PI Fishman JA;
XX
DR WPI, 1997-332804/30.
DR P-PDB; AAW32091-W32095.
XX
PT New nucleic acid from porcine retroviruses - used for detecting
PT viruses in transplant or other tissue and for assessing risk of
PT transmitting infection to graft recipient

```

```

XX
PS Claim 16; Fig 2; 128bp; English.
XX
CC This cDNA sequence represents a defective purified swine retrovirus
CC found in PK-15 cells containing the putative coding regions for viral
CC GAG, POL and ENV proteins. There are a few in frame stop codons and
CC apparent frame shifts in the given coding sequence which alter features
CC of the translation. This sequence and PCR fragments generated from the
CC sequence (see AAT74812-T74882) could be used to screen organs for the
CC presence of porcine retroviruses prior to xenotransplantation.
CC Transplantation can increase the likelihood of retroviral activation if
CC intact and infectious proviruses are present. The porcine retroviral
CC sequence can be used to generate probes to determine the level (e.g.
CC copy number) of intact (i.e. potentially replicating) porcine provirus
CC sequences in a strain of xenograft transplantation donors. It can be
CC used to detect mutations, genetic lesions or viral recombinants and to
CC determine the histological localisation of activated retroviruses. Using
CC Polymerase Chain Reaction DNA Quantitation (PQD) on blood mononuclear
CC cells, infectivity titration and susceptibility testing can be
CC performed. Ultimately animal donors without intact porcine retroviral
CC sequences or a lower copy number of viral elements could be selected.
XX
SQ Sequence 7393 BP; 2004 A; 1801 C; 1896 G; 1692 T; 0 other;

alignment_scores:
Quality: 2758.00 Length: 525
Ratio: 5.263 Gaps: 1
Percent Similarity: 99.810 Percent Identity: 99.619

alignment_block:
US-09-171-553B-4 x AAT74883 ..
Align seg 1/1 to: AAT74883 from: 1 to: 7393

1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisThrPh 17
|||||
598 ATGCGACAGACAGACTGACTACCCCTTAACTTGTGACTCTGACCATTTGAC 647
17 rglValArgSerArgAlaHisAsnLeuSerValGlnValLysGlyP 34
|||||
648 TGAAGTTAGATCCAGGCGCTCATTAATTGTGACGTGACGTTAAGAGGAC 697
34 rGTrpGlnThrPheCysAlaSerGluTrpProThrPheAspValGlyTrp 50
|||||
698 CTTCGACAGACTTTCTGTGCTCGAATGGCCCAACATTCATGTGATG 747
51 ProSerGluGlyThrPheAsnSerGluIleIleLeuAlaValLysAla 67
|||||
748 CCAATCAGAGGGGACCTTTAATTCTGAATTATCTGGCTGTTAAGGCANT 797
67 eilePheGlnThrGlyProGlySerHisProAspGlnGluProTyrIle 84
|||||
798 CATTTTTCAGACTGAGACCCGGGCTCATCCTGATCGAGGCCCTATATCC 847
84 eutThrTrpGlnAspLeuAlaGluAspProProTrpValLysProTrp 100
|||||
848 TTACGTGGCAGAGATTGGCAGAGATCTCCGCCATGAGGCTTAACCATG 897
101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGly 117
|||||
898 CTAAATTAACCAAGAAAGCCAGGTCCCGCAATCTTGCTTTGGAGAGAA 947
117 sAsnLysHisSerAlaGluLysValGluProSerSerSerTyrLeuPro 134
|||||
948 AAACAACAACCTGGCGGCAAAAAGTGGAGCCGC .TCCTGCTATCAACCC 996
134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150
|||||
997 GAGATCGAGAGAGCCGCCGACTTGGCCGGAACCCCAACCTGTTCCGCCACC 1046
151 ProLeuSerSerThrGlyCysGlyGlyThr.SerAlaProProGlyA 167
|||||

```

1047 CCTTATCCAGCAGAGGCTGTGTGAGGGGACCTCTGCCCCCTCTGGAG 1096
 167 LAPROVALValGluGlyProAlaAlaGlyThrArgSerArgArgGlyAla 183
 1097 CTCGGTGGTGGAGGAGCCTGCTGCGGAGCTCGAGCCGGAGAGGCGCC 1146
 184 ThrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTrp 200
 1147 ACCCCGAGCGGAGACAGAGATCGGATATTACCGCTCCGACCTATGG 1196
 200 yProPomMetProGlyGlyGlnLeuGlnProLeuGlnTrpProPhe 217
 1197 CCCCCTCATGCGAGGGGCCAATTCAGCCCTCCAGATTTGGCCCTTTT 1246
 217 eSerAlaAspLeuTrpAsnTrpLysThrAsnHisProProPheSerGlu 233
 1247 CTTTCGCAATCTCTATATTGGAAACTAACCATCCCCCTTTCTCGAG 1296
 234 AspProGluArgLeuThrGlyLeuValGluSerLeuMetPheSerHisG 250
 1297 GATCCCAACGCGCTCAGCGGCTGGTGAGTCCCTTATGTTCTCTACCA 1346
 250 nProThrTrpAspAspGlyGlnLeuLeuGlnThrLeuPheThrTrp 267
 1347 GCCTACTTGGAGATGTCTAACACAGCTCTGTCAGACACTCTTCACAC 1396
 267 LUGLARGGluArgIleLeuLeuGluAlaArgLysAsnValProGlyAla 283
 1397 AGGACGAGAGAGAAATTCGTAGAGGCTAAAAAAATGTTCTCTGGGCC 1446
 284 AspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPhePro 300
 1447 GACGGGCGACCCAGCAGTTCGCAAAATGAGATTGACATGGGATTTCC 1496
 300 uThrArgProGlyTrpAspTrpAsnThrAlaGluGlyArgGluSerLeu 317
 1497 GACTGCCCCGGTGGACTACACACCGCTGAAGGTAGGAGAGACTTGA 1546
 317 ySileTyArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArg 333
 1547 AAATCTATCGCCAGGCTCTGGTGGGGGCTCTCCGGGGCGCTCAAGAG 1596
 334 ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGlu 350
 1597 CCCACTAATTGGCTAAGGTAAAGAGAGGTGATGCAAGGACCGAAGAAC 1646
 350 OProSerValPheLeuGluArgLeuMetGluAlaPheArgArgPheThr 367
 1647 TCCCTGGGTATTCTTGTAGAGGCTCATGTGAAGCCCTTCAGGCGTTCA 1696
 367 rOpeAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 383
 1697 CTTTGTATCTTAACCTCAGAGGCCAGAAAGCTTCAGTGGCCCTGCTTC 1746
 384 IleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGlu 400
 1747 ATTTGGGAGTGGGCTCTGATATCAGGAAGAACTTCAGACATCGGAAG 1796
 400 yLeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluValT 417
 1797 GTTACAGAGGCTGAGTTACGTGATCTAGTGAGAGAGCAAGAGAGTGT 1846
 417 YTYTARGARGluThrGluGluGluLysGluGluArgLysGluLysGlu 433
 1847 ATTACAGAGGAGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1896
 434 ArgGluGluArgGluGluArgAspArgArgGlnGluLysAsnLeuTh 450
 1897 AGAGAAAG 1946
 450 rLysIleLeuAlaValAlaValGluGlyLysSerSerArgGluArgGlu 467
 1947 TAAAGTCTTGGCCCGAGTGGTTGAAGGAGAGAGAGAGAGAGAGAGAG 1996

467 rGAspPheArgLysIleArgSerGlyProArgGlnSerGlyAsnLeuGly 483
 1997 GAGATTTTGGAAAAATTAGTACAGCCCTTACAGAGTCAAGGAGACCTGGC 2046
 484 AsnArgThrProLeuAspLysAspGlnCysAlaTrpCysLysGlyLysG 500
 2047 AATAGAGACCCCACTCGACAAAGACAGTGTGCTATTGTAAAGAAAAAG 2096
 500 yHisTrpAlaArgAsnCysProLysLysGlyAsnLysGlyProLysVal 517
 2097 ACACGTGGGCAAGAGAACTGCCCAAGAGGAGAAACAAAGAGACCGAAG 2146
 517 euAlaLeuGluGluAspLysAsp 524
 2147 TAGCTCTAGAGAGATTAAGAT 2169
 seq_name: /sids2/gc9data/geneseq/geneseqn/NA2001.DAT: AAC67021
 seq_documentation block:
 ID AAC67021 standard; DNA; 6076 BP.
 AC AAC67021;
 DT 27-MAR-2001 (first entry)
 DE PERV env protein coding sequence SEQ ID NO: 21.
 KW Xenotransplantation; infectious agent; vaccine; ds.
 OS Porcine endogenous retrovirus.
 PN WO200071726-A1.
 PD 30-NOV-2000.
 PF 24-MAY-2000; 2000MO-US14296.
 PR 24-MAY-1999; 99US-0135631.
 PA (MAYO-) MAYO MEDICAL VENTURES.
 PI Federpiet MJ;
 DR WPI: 2001-032041/04.
 XX PT Inhibiting or preventing infectious agent transmission in mammalian
 PT transplant recipients, by introducing recombinant DNA comprising DNA
 PT encoding extracellular proteins of the agent into donor cells, such as
 PT swine cells -
 XX PS Claim 16; Page 107-109; 144pp; English.
 XX CC The present invention provides a method to prevent the transmission of
 CC infectious agents during xenotransplantation. This involves introducing
 CC to donor swine cells a recombinant DNA encoding a peptide fragment from
 CC the infectious agent, and then introducing these cells into the
 CC transplant recipient.
 XX SQ Sequence 6076 BP; 1613 A; 1512 C; 1609 G; 1342 T; 0 other;
 alignment_scores:
 Quality: 2659.00 Length: 525
 Ratio: 5.143 Gaps: 1
 Percent Similarity: 98.476 Percent Identity: 95.619
 alignment_block:
 US-09-171-553b-4 x AAC67021 ..
 Align seg 1/1 to: AAC67021 from: 1 to: 6076
 1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisTrpTh 17


```

1183 ATGGACGACGAGGTGACACCCCTTAGTTGACTCTCGACCTTGAC 1232
17  TGLuValArgSerArgAlaHisAsnLeuSerValGlnValLysLysLys 34
1233 TGAAGTTAAATCCAGGGCTCAATTTCTGCTAGTTACAGTTAAAGAGGAC 1282
34  ROTProGlnThrPheCysAlaSerGluThrProThrPheAspValGlyTyr 50
1283 CTTGGCCACACTTCTGTGTCTGTGAATGGCCGACATTGATGTGGATGG 1332
51  ProSerGluGlyThrPheAsnSerGluIleIleLeuAlaValLysAlaIle 67
1333 CCATCAGAGGGGACCTTTAATTCGAGATATCTCGCTGGCTTTAAAGCAAT 1382
67  eIlePheGlnThrGlyProGlySerHisProAspGlnGluProTyrIle 84
1383 TATTTTTCAGACTGACCCGCTCTCATCCCAATCAGAGCCCTATATCC 1432
84  euThrTrpGlnAspLeuAlaGluAspProProTyrProTyrValLysProTyr 100
1433 TTACGTGGCAAGATTGGCAGAGATCTCCGCCATGGGTTAAACCTTTGG 1482
101  LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyLysLys 117
1483 CTGAATTAAGCCCAAGAAAGCCAGTCCCGCAATTCGGCTTTGGAGAGAA 1532
117  sAsnLysHisSerAlaGlyLysValGluProSerSerSerTyrLeuPro 134
1533 AAACAAACACTCGCTGAAAGATCAAGCCCTC.TCTCATATCTATACCC 1581
134  rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150
1582 GAGATTGAGAGACCCCGCTGGCGGGAACCCCAATCTGTTCCGCCACC 1631
151  ProLeuSerSerThrGlyCysGlyLysLysThr.SerAlaProProGly 167
1632 CCTTATCTGGCACAGGTGCTGCGAGGGGACCCCTCCCTCGTAGAG 1681
167  laProValValGlnGlyProAlaAlaGlyThrArgSerArgArgLysAla 183
1682 CTCGGCGGTGGAGGACCTGCTGCAAGGACTCGGAGCGGAGGGGCCCC 1731
184  ThrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrGln 200
1732 ACCCGGAGCGGACAGACAGATTCGCAATTAACCGTCGCGAGCTACGG 1781
200  yProMetProGlyGlyGlnLeuGlnProLeuGlnTyrTyrProPheS 217
1782 CCGTCCCAACCGGGGGCCAAATTCACCCCTCCAGTATTCGCCCTTTT 1831
217  eSerAlaAspLeuTyrAsnTrpLysThrAsnHisProProPheSerGlu 233
1832 CTTTGCACATCTCTATAATTGAAAACTAACCAATCCCCCTTTTCGGAG 1881
234  AspProGlnArgLeuThrGlyLeuValGlnSerLeuMetPheSerHisGln 250
1882 GATCCCAACCGCTCACGGGGTGTGGTAGTCCCTTATGTTCTCTACCA 1931
250  nProThrTrpAspAspCysGlnGlnLeuGlnThrLeuPheThrTrpG 267
1932 GCCTACTGGGATGATTCACACACTCTGACAGACCTTCACAAACCG 1981
267  luGluArgGluArgIleLeuLeuGluAlaArgLysAsnValProGlyAla 283
1982 AGGAGCGAGAGAAATCTGTAGAGCTAGAAAAAATGTTCTCTGGGCC 2031
284  AspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPhePro 300
2032 GACGGCGGACCCAGCTGAGTGAATAATGATGATGAGATGGATTTCCCT 2081
300  uThrArgProGlyTrpAspTyrAsnThrAlaGluGlyArgGluSerLeu 317
|||||
2082 GACTCGCCCGGTTGGACTACAAACAGCGGCTGAAGGTAGAGAGCTTGA 2131
317  ystIleTyrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgTyr 333
2132 AAATCTATGCGCCAGGCTGTGTGGCGGCTTCGGGGCCCTCCAAAGACGG 2181
334  ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGluPr 350
2182 CCCACTAATTTGGCTAAGGTAAAGAGAGTATGACAGGACCCGAATGAAC 2231
350  oProSerValPheLeuGluArgLeuMetGluAlaPheArgTyrPheThr 367
2232 CCCCTCTGTTTCTCTGAGAGGCTCTGGAAGCCCTTCAGGCGGTACACC 2281
367  rPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 383
2282 CTTTGTATCCCACTCAGAGGCCCAAAAGCCCTCAATGGCTTTGGCTTT 2331
384  IleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGlu 400
2332 ATAGGACAGTACAGCTTGGATATTAGAAAGAGCTTCAGAGACTGGAGG 2381
400  yLeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGlyValTyr 417
2382 GTTACAGAGAGCTGAGTTACGTATCTAGTGAAGAGGCGAGAAAGTAT 2431
417  yTyrArgArgGluThrGluGluGlyLysGlnGlnArgLysGlyLysGlu 433
2432 ATTACAAAGAGGACAGAAAGAAAGGAAACCAAAAGAAAGAGAGAA 2481
434  ArgGluGluArgGluGluArgArgAspArgArgGlnGlyLysAsnLeu 450
2482 AGAGAGAAAGGAGGAAAGACCTAATAAACGCAAGAGAAAGAAATTGAC 2531
450  rLysIleLeuAlaAlaValAlaGluGlyLysSerSerArgGluArgGlu 467
2532 TAAAGATTGTGCTGCAAGTGTGAAGAGAAACCAATACGAAAGAGAA 2581
467  rGAspPheArgLysIleArgSerGlyProArgGlnSerGlyAsnLeuGly 483
2582 GAGATTTTAGAAATTAAGTACAGGCCCTTAGACAGTACAGGAACTGGGC 2631
484  AsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGln 500
2632 AATAGACCCCACTCGACAGAGCAATGTGCTATTTGTAAGAAAAAGG 2681
500  yHisTrpAlaArgAsnCysProLysLysGlyAsnLysGlyProLysVal 517
2682 ACACGTGGCAAGGACAGTCCCAAGAAAGGAAACAAAGAGACTGAAGTCT 2731
517  euAlaLeuGluGluAspLysAsp 524
2732 TAGCTCTGAAAGAAATTAAGAC 2754
seq_name: /STD2/gcgdata/geneseq/geneseqn/NA2001.DAT:AMC67032
seq_documentation_block:
ID AAC67032 standard; DNA; 6076 BP.
XX
AC AAC67032;
XX
DT 27-MAR-2001 (first entry)
XX
DE PERV env protein coding sequence SEQ ID NO: 32.
XX
KM Xenotransplantation; infectious agent; vaccine; ds.
XX
OS Porcine endogenous retrovirus.
XX
PN MO200071726-A1.
XX
PD 30-NOV-2000.
XX

```

PF 24-MAY-2000; 2000MO-US14296.
 XX
 XX 24-MAY-1999; 99US-0135631.
 XX
 PA (MAYO-) MAYO MEDICAL VENTURES.
 XX
 PI Federspiel MJ;
 XX
 DR WPI; 2001-032041/04.
 XX
 XX Inhibiting or preventing infectious agent transmission in mammalian
 PT transplant recipients, by introducing recombinant DNA comprising DNA
 PT encoding extracellular proteins of the agent into donor cells, such as
 PT swine cells -
 XX
 PS Claim 16; Page 117-119; 144pp: English.
 XX
 CC The present invention provides a method to prevent the transmission of
 CC infectious agents during xenotransplantation. This involves introducing
 CC to donor swine cells a recombinant DNA encoding a peptide fragment from
 CC the infectious agent, and then introducing these cells into the
 CC transplant recipient.
 XX
 SQ Sequence 6076 BP; 1619 A; 1505 C; 1598 G; 1354 T; 0 other;
 alignment_scores:
 Quality: 2659.00 Length: 525
 Ratio: 5.143 Gaps: 1
 Percent Similarity: 98.476 Percent Identity: 95.619
 alignment_block:
 US-09-171-553b-4 x AAC67032 ..
 Align seg 1/1 to: AAC67032 from: 1 to: 6076
 1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrp 17
 1183 ATGGGACAGACGGGTGACGACCCCTCTAGTTGACTCGACCATTTGGAC 1232
 17 rclLValArgSerArgAlaHisAsnLeuSerValGlnValLysGlyP 34
 1233 TGAAGTTAAATCCAGGGCTCATTAATTTGCAAGTTCAGGTTAAGAGGAC 1282
 34 rdtGlnThrPheCysAlaSerGluTrpProThrPheAspValGlyTrp 50
 1283 CTGGGACGCTTTCTGTCTCTGATGAGCCGACATTCGATGTTGGATGG 1332
 51 ProSerGluGlyThrPheAsnSerGluIleIleLeuAlaValLysAla 67
 1333 CCATCAGAGGGGACCTTAAATTCAGATTAATCCCTGGCTGTTAAAGCAAT 1382
 67 ellePheGlnThrGlyProGlySerHisProAspGlnLupProTyrIle 84
 1383 TATTTTTCAGACTGAGACCCGGCTCTCATCCCATCAGAGCCCTATATCC 1432
 84 eutThrTrigLAspleuAlaGluuAspProProProTyrValLysProTrp 100
 1433 TTACGTGCGACATTTGGCAGAGATCCTCCGCCATGGGTTAAACCTTGG 1482
 101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGlu 117
 1483 CTGATTAAGCCAGAAAGCCAGGTCCCGCATTTCTGGCTCTGGAGAGAA 1532
 117 sAsnLysHisSerAlaGluLysValGluProSerSerSerTyrLeuPro 134
 1533 AAACAAACACTCGCTGAAAAAGTCACACCCCTC; TCCTCATATCTACCCC 1581
 134 rGAspArgGlyAlaAlaAspleuAlaGluuAspProProThrCysSerProThr 150
 1582 GAGATTTGAGGAGCCCGCTTGGCGGAGACCCCAATCTGTTCCCCACCC 1631
 151 ProLeuSerSerThrGlyCysCysGluGlyThrSerAlaProProGly 167

1632 CCCTATCTGGACAGGGGTGCTGGAGGGAGACCTCTCCCTCTGGAG 1681
 167 laProValValGluGlyProAlaAlaGlyThrArgSerArgArgGlyAla 183
 1682 CTCGGCGGTGGAGGAGCTCTGGAGGACTCGGAGCGGAGGAGGGCC 1731
 184 ThrProGluArgThrAspGluIleAlaIleLeuProLeuAlaGlyTrpGly 200
 1732 ACCCGGAGCGAGACAGAGATCGCACATTTACCGCTGCGACGTACGG 1781
 200 yProProMetProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPhe 217
 1782 CCCTCCACACCGGGGGGCCAATTGACGCCCTCCAGATTTGGCCCTTTT 1831
 217 erSerAlaAspleuTyrAsnTrpLysThrAsnHisProProPheSerGlu 233
 1832 CTTCGTCAGATCTCTATATTGAAACATACTACCATCCCTTTCTCGGAG 1881
 234 AsproGlnArgLeuThrGlyLeuValGluSerLeuMetPheSerHisGly 250
 1882 GATCCCAACGCTTCACGGGGTGTGGAGTCCCTTATGTTCTCTCACCA 1931
 250 nProThrTrpAspAspCysGlnGlnLeuLeuGlnThrLeuPheThrG 267
 1932 GCCTACTTGGGATGATTTGTCAACAGCTGCTGACAGACCTCTTCAACACG 1981
 267 lGluArgGluArgIleLeuLeuGlnAlaArgLysAsnValProGlyAla 283
 1982 AGGAGCGAGAGAGATTCGTGTAGAGGCTAGAAAAATGTTCCCTGGGGCC 2031
 284 AspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLe 300
 2032 GACGGGGACCCACGAGTTGCAAAATGAGATTGACATGAGGATTTCCCTT 2081
 300 uThrArgProGlyTrpAspTrpAsnThrAlaGluArgGluSerLeu 317
 2082 GACTCGCCCGGCTTGGAGCTTCAACACGCTGCTGAGAGGAGGAGGCTTGA 2131
 317 ySileTyrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArg 333
 2132 AAATCTATTCGACAGGCTCTGTGGCGGGTCTCCGGGGCGCTCAAGACGG 2181
 334 ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGluP 350
 2182 CCCACTAATTTGGCTAAGTAAGAGAAAGTATGACGAGGACCGAATGAC 2231
 350 oProSerValPheLeuGluArgLeuMetGluAlaPheArgArgPheThrP 367
 2232 CCCTCTGTGTTTCTTGAAGGGCTCTTGAAGCCTTGAAGCGGTACACCC 2281
 367 rPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 383
 2282 CTTTGTATCCACCTCTGAGAGGCCCAAAAGCCTCAGTGGCTTTGGCTTT 2331
 384 IleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGlu 400
 2332 ATAGACAGCTCAGGCTTGATATTAGAAAGAACCTTCAGACACTGGAAG 2381
 400 yLeuGlnGluAlaGluLeuArgAspleuValArgGluAlaGluLysVal 417
 2382 GTTACAGAGAGGCTGAGTTACGTGATCTAGTGAAGGAGGAGAGAAAGTAT 2431
 417 yTyrArgArgGluThrGluGluLysGluGlnArgLysGluLysGlu 433
 2432 ATTACAAAGAGGAGAGAGAGAAAGAAAGGACAAAGAAAGAGAGAGAA 2481
 434 ArgGluGluArgGluGluArgAspArgArgGlnGluLysAsnLeuThr 450
 2482 AGAGAGAGAAAGGAGAGAAAGCATAATTAACGCAAGAGAAAGATTTGAC 2531
 450 rLysIleLeuAlaValValGluGlyLysSerSerArgGluArgGlu 467

2532 TAAGATCTTGCTGCTGAGTGTGAAGGGAACAAATACGGAAGAGA 2581
 467 rGAspPharGlyLysIleargSerGlyProargGlnSerGlyAsnLeuGly 483
 2582 GAGATTTTGAAGAAATTAGTCAAGCCCTAGACAGTCAAGGAGACCTGGGC 2631
 484 AsnArgTrpProLeuAspLysAspGlnCysAlaTrpCysLysGluLysG 500
 2632 AATAGACCCCACTGACAGACCAATGTCATATTTAAAGAAAAAG 2681
 500 yHISrPalargAsnCysProLysGlyAsnLysGlyProLysVal 517
 2682 ACACGGGCAAGGAGACTGCCCAAGAAAGAAAGAAAGAAAGAGTCT 2731
 517 euAlaLeuGluGluAspLysasp 524
 2732 TAGCTCTGGAAGAGATTAAGAC 2754
 seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT: AAC67023

seq_documentation_block:

ID AAC67023 standard; DNA: 7873 BP.

AC AAC67023:

DT 27-MAR-2001 (first entry)

DE PERV env protein coding sequence SEQ ID NO: 23.

KW Xenotransplantation; infectious agent; vaccine; ds.

OS Porcine endogenous retrovirus.

PN W0200071726-A1.

PD 30-NOV-2000.

PE 24-MAY-2000; 2000WO-US14296.

PR 24-MAY-1999; 99US-0135631.

PA (MAYO-) MAYO MEDICAL VENTURES.

PI Federspiel MJ;

DR WPI; 2001-032041/04.

PT Inhibiting or preventing infectious agent transmission in mammalian

PT transplant recipients, by introducing recombinant DNA comprising DNA

PT encoding extracellular proteins of the agent into donor cells, such as

PT swine cells -

PS Claim 16; Page 112-115; 144pp; English.

CC The present invention provides a method to prevent the transmission of

CC infectious agents during xenotransplantation. This involves introducing

CC to donor swine cells a recombinant DNA encoding a peptide fragment from

CC the infectious agent, and then introducing these cells into the

CC transplant recipient.

SO Sequence 7873 BP; 2200 A; 1913 C; 1962 G; 1798 T; 0 other;

alignment_scores:

Quality: 2659.00

Ratio: 5.143

Percent Similarity: 98.476

Length: 525

Gaps: 1

Percent Identity: 95.619

alignment_block:

us-09-171-553b-4 x AAC67023

Align seg 1/1 to: AAC67023 from: 1 to: 7873

1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisTrpTh 17
 62 ATGGAGACAGACGGGTGACAGACCCCTTACTTGTACTCTGACCATTTGAC 111
 17 rGluValArgSerArgAlaHisAsnLeuSerValGluValLysGlyP 34
 112 TGAAGTAAATCCAGGCTCATATATTGTCAAGTTCAGGTTAAAGAAAGGAC 161
 34 rGluValArgSerArgAlaHisAsnLeuSerValGluValLysGlyP 50
 162 CTGGGAGAGATTTCTGTGTCTGATGAGGCGGACATTCATGTTGATGATG 211
 51 ProSerGluGlyThrPheAsnSerGluIleIleLeuAlaValLysAla 67
 212 CCATCAGAGGGGACCTTAATTCGAGATTATCTGCTGCTGCTGTTAAACAAAT 261
 67 eLlIephGlnThrGlyProGlySerHisProAspGlnGluProTyrIle 84
 262 TATTTTTCAGACTGGACCGGCTCTCATCCCAATTCAGAGCCCTATATTC 311
 84 eutHrTrpGlnAspLeuAlaGluAspProProProTrpValLysProTrp 100
 312 TTACGTGCGCAAGATTTGGCAGAGGATCTCCGCGCATGGGTTAAACCTTGG 361
 101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGly 117
 362 CTGAATTAAGCCAAAGAAAGCCAGGTCGCCGAATTCCTGCTTGGAGAGAA 411
 117 sAsnLysHisSerAlaGluLysValGluProSerSerSerTyrLeuPro 134
 412 AAACAAACACTCGGCTGAAAAAGTCAAGCCCTC.TCCTCATATCTACCC 460
 134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150
 461 GAGATTGAGAGCCCGCGGCTTGCGGAAACCCCAATCTGTCGCCACCC 510
 151 ProLeuSerSerThrGlyCysCysGlyGlyThr.SerAlaProProGly 167
 511 CCTTATCTGCGCACAGGCTGCTGCGAGGGAGCCCTGCTCCCTCTGAG 560
 167 lApProValGluGlyProAlaAlaGlyThrArgSerArgArgGlyAla 183
 561 CTCGGGCGGTGAGAGACTGCTGCGAGGACTGGAGCCGAGGGGCGCC 610
 184 ThrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyr 200
 611 ACCCGGAGCGAGACAGAGATCGGACATTAACCTGCGGACGACG 660
 200 yProProMetProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPhe 217
 661 CCCCTCCACACCGGGGGCCCAATTTGACGCCCTCCAGTATTTGGCCCTTT 710
 217 eSerAlaAspLeuTyrAsnTrpLysThrAsnHisProProPheSerGlu 233
 711 CTTTGCAGATCTCTATATTTGGAACCTAACCATCCCTTTCTGCGAG 760
 234 AspProGlnArgLeuThrGlyLeuValGluSerLeuMetPheSerHis 250
 761 GATCCCAACAGCCCTCACGGGGTTGTGTGAGTCCCTATGTTCTCTACCA 810
 250 nProThrTrpAspAspCysGlnGlnLeuLeuGlnThrLeuPheThrTh 267
 811 GCCCTACTTGGATGATGATGTCACACGCTGCTGACACACCTCTTCAACAC 860
 267 lGluArgGluArgIleLeuLeuGlnAlaArgLysAsnValProGlyAla 283
 861 AGAGCGAGAGAGATTTCTGTAGAGCTGAGAAAAATGTTCTCGGGGCC 910
 284 AspGlyArgProThrGlnLeuGlnAsnGluIleAspMetClyPhePro 300
 911 GACGGCGAGCCACGACGATTGCAAAATGAGATTGACATGAGGATTTCCCT 960
 300 uThrArgProGlyTrpAspTyrAsnThrAlaGluGlyArgGluSerLeu 317

XX	24-MAY-2000; 2000MO-US14296.
PF	
XX	
PR	24-MAY-1999; 990S-0135631.
XX	
XX	
PA	(MAYO-) MAYO MEDICAL VENTURES.
XX	
PI	Federspiel MJ;
DR	
XX	WPI; 2001-032041/04.
XX	
PT	Inhibiting or preventing infectious agent transmission in mammalian
PT	transplant recipients, by introducing recombinant DNA comprising DNA
PT	encoding extracellular proteins of the agent into donor cells, such as
PT	swine cells -
XX	
PS	Claim 16; Page 101-104; 144pp; English.
XX	
CC	The present invention provides a method to prevent the transmission of
CC	infectious agents during xenotransplantation. This involves introducing
CC	to donor swine cells a recombinant DNA encoding a peptide fragment from
CC	the infectious agent, and then introducing these cells into the
CC	transplant recipient.
XX	
SO	Sequence 7362 BP; 1997 A; 1821 C; 1663 T; 0 other;

[illegible]

151	ProLeSerSerThrGlyCysCysGlnGlyThr.SerAlaProProGly	167
1697	CCCCATATCTGGAGACAGGGTGGCGAGGGGACCCCTCTGCCCTCTGGAG	1746
167	LaProValValGluGlyProAlaAlaGlyThrArgSerArgArgGlyAla	183
1747	CTCCGGCGGTGGAGAGCACTGCTGCAGGAGCACTCGAGCGGAGGAGGGGCC	1796
184	ThrProGluArgThrAspGluLeuAlaIleLeuProLeuArgThrArg	200
1797	ACCCTGGAGCGGACAGACAGATCGCAGATTACCGCTGGCCACGTCACGG	1846
200	YProProMetProGlyGlyGlnLeuGlnProLeuGlnTrpTrpProPheS	217
1847	CCCTCCCAACCGGGGGGCCAATTTCACACCCCTCCAGTATGGCCCTTTT	1896
217	erSerAlaAspLeuTrpAsnTrpIysThrAsnHisProProPheSerGlu	233
1897	CTTCTGCAAGATCTCTATATTTGAAACATAACCAATCCCTCTTCTCGAG	1946
234	AspProGlnArgLeuThrGlyLeuValGluIserLeuMetPheSerHisGlu	250
1947	GATCCCCAACGCTCACGGGGTGGAGTCCCTTATGTTCTCTCACCA	1996
250	nProThrTrpAspAspCysGlnGlnLeuLeuGlnThrLeuPheThrArg	267
1997	GCCTACTTGATGGATGATGGCAACACACTGCTCGACACACTTTCACACCG	2046
267	GluGluArgGluArgIleLeuLeuGluAlaArgIysAsnValProGlyAla	283
2047	AGGAGGAGAGAGAAATTCGTTAGAGGCTAGAGAAAAATGTTCTCGGGCC	2096
284	AspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLe	300
2097	GACGGCGACCAACCGACGTGCAAAATAGATTGACATGGGATTTCCCTT	2146
300	uThrArgProGlyTrpAspTrpAsnThrAlaGluGlyArgGluSerLeu	317
2147	GACTCGCCCCGGTTGGGACTACACACGCTGAAAGTGGGAGAGCTTGA	2196
317	ysIleLeuArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArg	333
2197	MAATCATGCGCCAGGCTCTGCTGGCGGGCTTCGGGGGCCCTCAAGACGG	2246
334	ProThrAsnLeuAlaIysValArgGluValMetGlnIyProAsnGluPr	350
2247	CCACATAATTGGCTAAAGGTAGAGAACTGATGACGGGACCGAATGAAAC	2296
350	oProSerValPheLeuGluArgLeuMetGlnAlaPheArgArgPheThrP	367
2297	CCCTCTGTTTTCTTGGAGAGCTCTTGGAAACCTTCAAGCGGTACACCC	2346
367	roPheAspProThrSerGluAlaGlnIysAlaSerValAlaLeuAlaPhe	383
2347	CTTTGATCCCACTCAGAGGCCCAAAAAAGCTCAGTGGCTTGGCCCTTT	2396
384	IleGlyGlnSerAlaLeuAspIleArgIysIysLeuGlnArgLeuGlu	400
2397	ATAGGACATCGACCTTGATATTGTGAAAGAGCTTCAGAGCTCGAAGG	2446
400	yleuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluIysValT	417
2447	GTTACAGGAGCGCTGAATTTCGATCTGATGTAAGAGGACAGAGAAATAT	2496
417	yTrpTrpArgArgGluThrGluGluGluIysGluGlnArgIysGluIysGlu	433
2497	ATTACCAAAAGGGGACAGAGAAAGAAAGGAGCAAAAGAAAGAGAGAA	2546
434	ArgGluGluArgGluGluArgArgAspArgArgGlnGluIysAsnLeuThr	450
2547	AGGAGGAAAGGAGGAGAAAGACTTAATAACGCAAGAGAAAGAAATTTGAC	2596
450	rIysIleLeuAlaAlaValValGluGlyLysSerSerArgGluArgGlu	467

```

XX          2597  TAAGATCTTGCTGCACAGTGTTGAAGGAAAGCAATATACGAAAGAGACA 2646
XX          |||||||.....:|||||
XX          467  rgaSPpHeaRlySLlearySerclYProaRgInSerGlyaSnleagUly 483
XX          |||||||.....:|||||
XX          2647  GAGATTTTATAGAAAATTAGGTCAAGCCCTGACACAGTCAGGGAACTGGGC 2696
XX          |||||||.....:|||||
XX          484  AsnaRqThrProLeuAspLySASPGLnCysAlaTyICysLySgLuLySgI 500
XX          |||||||.....:|||||
XX          2697  AATAGAGACCCACCTCGAACAGGACCATGTGCATATTGTAAAGAAAAAGG 2746
XX          |||||||.....:|||||
XX          500  yHISrPAlaArGAsnCysProLySLySGlyAsnLySGlyProLySValL 517
XX          |||||||.....:|||||
XX          2747  ACACGTGGGCAAGAACTGCCCAAGGAAGGAAACAAAGACTGAGAGCTT 2796
XX          |||||||.....:|||||
XX          517  euAlaLeuGLuGLuAspLySASP 524
XX          |||||||.....:|||||
XX          2797  TAGCTCTGAAAGAAATTAAGAC 2819
XX          |||||||.....:|||||
XX          seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:AAIT74884
XX          seq_documentation_block:
XX          ID AAT74884 standard; cDNA; 7892 BP.
XX          AAT74884:
XX          AC
XX          09-FEB-1998 (first entry)
XX          DE Miniature swine retrovirus CDNA.
XX          KM Retrovirus; porcine; GAG protein; POL protein; ENV protein;
XX          KM xenotransplantation; Infectious; provirus; organ transplant; donor.
XX          KW activated virus; PCR; ss.
XX          OS Porcine retrovirus.
XX          FH Location/Qualifiers
XX          FT CDS 585..2155
XX          FT /*tag= a
XX          FT 585..2156
XX          FT mat_peptide
XX          FT /*tag= b
XX          FT /note= "putative GAG protein"
XX          FT CDS 2307..5744
XX          FT /*tag= c
XX          FT mat_peptide 2307..5741
XX          FT /*tag= d
XX          FT /note= "putative POL protein"
XX          FT CDS 5620..7536
XX          FT /*tag= e
XX          FT mat_peptide 5620..7533
XX          FT /*tag= f
XX          FT /note= "putative ENV protein"
XX          PN WO9721836-A1.
XX          PN 19-JUN-1997.
XX          PD 13-DEC-1996; 96WO-US19680.
XX          PR 14-DEC-1995; 95US-0572645.
XX          PA (GEHO ) GEN HOSPITAL CORP.
XX          PI Fishman JA;
XX          DR WPI: 1997-332804/30.
XX          DR P-PSDB; AAM32096-W32098.
XX          New nucleic acid from porcine retro:viruses - used for detecting
XX          PT viruses in transplant or other tissue and for assessing risk of
XX          PT transmitting infection to graft recipient
XX          CS Claim 22: Fig 3; 128bp: English.

```

XX This cDNA sequence represents a porcine retrovirus from miniature swine
 CC containing the putative coding regions for viral GAG, POL and ENV
 CC proteins. This sequence and PCR fragments generated from it
 CC (see AAT74812-T74882) can be used to screen organs for the presence of
 CC porcine retroviruses prior to xenotransplantation. Transplantation can
 CC increase the likelihood of retroviral activation if intact and
 CC infectious proviruses are present. The porcine retroviral sequence can be
 CC used to generate probes to determine the level (e.g. copy number) of
 CC intact (i.e. potentially replicating) porcine provirus sequences in a
 CC strain of xenograft transplantation donors. It can be used to detect
 CC mutations, genetic lesions or viral recombinants and also to determine
 CC the histological localisation of activated retroviruses. Using Polymerase
 CC Chain Reaction DNA quantitation (PQD) on blood mononuclear cells,
 CC infectivity titration and susceptibility testing can be performed.
 CC Ultimately animal donors without intact porcine retroviral sequences or a
 CC lower copy number of viral elements could be selected.
 XX
 SQ Sequence 7892 BP; 2191 A; 1915 C; 1980 G; 1806 T; 0 other;

alignment_scores:

Quality: 2640.00 Length: 525
 Ratio: 5.116 Gaps: 1
 Percent Similarity: 98.286 Percent Identity: 94.476

alignment_block:

US-09-171-553b-4 x AAT74884

Align seg 1/1 to: AAT74884 from: 1 to: 7892

1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisThrTrp 17
 585 ATGGGACAGACGGTGTGACACCCCTTAGTTGACCTGACCATTTGGAC 634
 17 rgluValArgSerArgAlaHisLeuSerValGlnValLysLysGly 34
 635 TANAATTAAATCCAGGGGCTCAATTTTGACATTGAGTTAAGAGGAG 684
 34 roTrpGlnThrPheCysAlaSerGluTrpProThrPheAspValGlyTrp 50
 685 CTGGCAGACTTCTGTGTCTGATGAGGCGACATTCGATGCTTGATGG 734
 51 ProSerGluGlyThrPheAsnSerGluLeuLeuAlaValLysAla 67
 735 CCATAGAGGGGACCTTAATTCAGATTTATCCGCTGTTAAAGCAGT 784
 67 eilePheGlnThrGlyProGlySerHisProAspGlnGluProTyrIle 84
 785 TATTTTTCAGACTGAGACCGGCTCTCATTCGCCGATCAGGAGCCATATATCC 834
 84 euThrTrpGlnAspLeuAlaGluAspProProProTrpValLysProTrp 100
 835 TTACGTGGCAAGATTGGCAGAGATCTCCGCCATGGGTTAAACCAAGG 884
 101 LeuAsnLysProArgLysProGlyProArgGlleuAlaLeuGlyGly 117
 885 CTGAATTAAGCCAAAGAAAGCCAGGTCGCCGAATTCGGCTTGGAGAGAA 934
 117 AsnLysHisSerAlaGluLysValGluProSerSerSerLysLeuPro 134
 935 AAACAAACACTCGCTGAAAAAGTCAGCCCTC, TCCCATATATCTACCC 983
 134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150
 984 GAGATTGAGAGACCAACCGCTTGGCCGGAACCCCAATCTGTTCGCCAC 1033
 151 ProLeuSerSerThrGlyCysCysGlyGlyThrSerAlaProProGly 167
 1034 CCTTATCTGGACAGGGGTGGCCGAGGAGACCTTTGCCCTCTCTGAG 1083
 167 laProValValGluGlyProAlaAlaGlyThrArgSerArgArgGlyAla 183

1084 CTCGGCGGTGGAGGACCTGCTGCAGGGACTCGAGGCCGGAAGGCGCC 1133
 184 ThrProGluAlaArgThrAspGluIleAlaIleLeuProLeuArgThrTrp 200
 1134 ACCCGGAGCGGACAGACGAGATCGGACATTACCGCTGGCAGCTACG 1183
 200 yProProMetProGlyLysGlnLeuGlnProLeuGlnIleThrTrpPhe 217
 1184 CCTCCACACCGGGGGGCAATTCAGCCCTCCAGATTGGCCCTTTT 1233
 217 eSerAlaAspLeuThrAsnTrpLysThrAsnHisProProPheSerGlu 233
 1234 CTTTCGAGATCTATATATGAAAACTAACATCCCTCTCTCGGAG 1283
 234 AsProGlnAlaGluThrGlyLeuValGlnSerLeuMetPheSerHisG 250
 1284 GATCCCAACGCCCTCACGGGTTGGTGAGCCCTTATGTTCTCTCACA 1333
 250 nProThrTrpAspCysGlnGlnLeuLeuGlnThrLeuPheThrTrp 267
 1334 GCCTACTGGGATGATTGTCAACAGCTGCTCAGACACTTTCACAAACG 1383
 267 lGluArgGluArgIleLeuLeuGluAlaArgLysAsnValProGlyAla 283
 1384 AGGACCGAGAGAGATTTCTATTAGGCTACAAAAATGTTCCGCGGCG 1433
 284 AspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPhePro 300
 1434 GACGGGCACCCACCGCGTTGCAAAATGAGATTGACATGGATTTCCTT 1483
 300 uThrArgProGlyTrpAspTrpAsnThrAlaGlnGlyArgGlnSerLeu 317
 1484 AACTCGCCCGGTTGGGACTACAAACAGCTCGAAGGTAAGGAGGCTGA 1533
 317 yIleThrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArg 333
 1534 AAATCTATCGCCAGGCTCTGTTGGGCTCCGCGGCGCTCAAGACGG 1583
 334 ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGly 350
 1584 CCCCTAAATTCCTAAGGTAAAGTAAGTGAAGGACCGAATGAAC 1633
 350 oProSerValPheLeuGlnArgLeuMetGluAlaPheArgArgPheThr 367
 1634 CCCCTGTTTCTTCTGAGAGGCTCTGGAAGCTTCAGCGGATACACC 1683
 367 roPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 383
 1684 CTTTGTATCCCACTCAGAGGCCCAAAAGCCCTCAGTGGCTTGGCCTT 1733
 384 IleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGlu 400
 1734 ATAGGACAGTCAAGCTTGGATATTAGAAAGAACTTCAGAGCTGGAAG 1783
 400 yLeuGlnGluAlaGluLeuAspLeuValArgGluAlaGlyValTrp 417
 1784 GTTACAGAGAGGTGAGTTAGCTATCATGCAAGGAGCGAGAAAGTAT 1833
 417 yTrpArgArgGluThrGlnGlnGlnLysGlnGlnArgLysGlyLysGlu 433
 1834 ATTTACAAAAGGAGAGACGAAGAAAGGAAAGGAAAGGAAAGAGAGAA 1883
 434 ArgGlnGluArgGluGluArgArgAspArgArgGlnGlnLysAsnLeu 450
 1884 AGAGAGAAAGGAGAGAAAGCATAATTAACGCAAGAGAAATTTGAC 1933
 450 rLysIleLeuAlaAlaValValGlnGlyLysSerSerArgArgGluAla 467
 1934 TAAGATCTTGCTGCAAGTGGTTGAAGGAAAGCAATAGGAAAGAGAA 1983
 467 rGAspPheArgLysIleArgSerGlyProArgGlnSerLysAsnLeuGly 483
 1984 GAGATTTTAAAGAAATTAAGTCAAGGCCCTTAAGACAGTCAAGGAAAC 2033

484 AsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGln 500
 |||||
 2034 AATAGACCCCTCGACAGACCAATGTCATATGTAAGAAAGAGG 2083
 500 YHISTPAlaArgAsnCysProLysLysGlnLysGlnLysValI 517
 |||||
 2084 ACACCTGGGCAAGAACTGCCCCAAGAGGAAACAAAGCACCAAGATCC 2133
 517 euAlaLeuGluGluAspLysAsp 524
 |||||
 2134 TAGCTCTAGAGAAAGATAAGAT 2156
 seq_name: /SID52/gcseqdata/geneseq/geneseqn/AA2001.DAT:AAF77727
 seq_documentation_block:
 ID AAF77727 standard; DNA; 8132 BP.
 XX
 AC AAF77727;
 XX
 XX 23-MAY-2001 (first entry)
 DE Nucleotide sequence of a retrovirus found in miniature swine.
 XX
 XX Retrovirus; graft transplantation; xenotransplantation; miniature swine;
 KM SS.
 XX
 XX unidentified.
 OS
 XX
 PN US6190861-B1.
 XX
 PD 20-FEB-2001.
 XX
 XX 13-DEC-1996; 96US-0766528.
 PF
 XX 14-DEC-1995; 95US-0572645.
 PR
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 DR WPI: 2001-256211/26.
 DR P-PSDB; AAB73285, AAB73286, AAB73287.
 XX
 PT Assessing risk of endogenous retroviruses in clinical practice and in
 PT xenotransplantation, comprises using probe sequences derived from swine
 PT or miniature swine retroviral genome.
 PS
 PS Claim 1; Fig 3; 127pp; English.
 XX
 CC The present invention relates to a method for screening a cell or tissue
 CC for the presence or expression of a retrovirus (RV), comprising
 CC contacting a target nucleic acid from the cell or tissue with a second
 CC nucleic acid from the present invention (e.g. the present sequence or a
 CC fragment thereof). The method is useful for RV detection and to assess
 CC graft transplantation risk. Screening of animals allows the elimination
 CC of donors with active replication of known viruses. Inactive proviruses
 CC can be detected and inactivated, allowing identification and elimination
 CC of potential human pathogens derived from swine in a manner not possible
 CC in the outbred human organ donor population and is important to the
 CC development of human xenotransplantation.
 CC
 SO Sequence 8132 BP; 2248 A; 1977 C; 2037 G; 1870 T; 0 other;

alignment_scores:
 Quality: 2640.00 Length: 525
 Ratio: 5.116 Gaps: 1
 Percent Similarity: 98.286 Percent Identity: 94.476

alignment_block:
 US-09-171-553b-4 x AAF77727 ..

Align seg 1/1 to: AAF77727 from: 1 to: 8132

1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrpTh 17
 |||||
 585 ATGGACAGACGCGTACGACCCCTTATGTTGACTCTCGACCATTTGAC 634
 17 rGluValArgSerArgAlaHisAsnLeuSerValGlnValLysLysGlyP 34
 |||||
 635 TGAAGTTAATATCCAGGGCTCATATTGTGATTCAGTTAGTAAGAGGAC 684
 34 rotRPolnThrPheCysAlaSerGluTyrProThrPheAspValGlyTyr 50
 |||||
 685 CTGGGACAGCTTCTGTGTCTCTGATGCGCACATTCGATGTTGATGG 734
 51 ProSerGluGlyThrPheAsnSerGluLeuLeuAlaValLysAlaI 67
 |||||
 735 CCATCGAGGGAGCCTTATTTCTGACATTTATTCGGCTGTTAAAGCAT 784
 67 eilePheGlnThrGlyProGlySerHisProAspGlnGluProTyrIle 84
 |||||
 785 TATTTTTCAGACTGGACCGGCTCATCCGATCCGATCAGAGCCCTATCC 834
 84 euThrTrpGlnAspLeuAlaGluAspProProProTyrValLysProThr 100
 |||||
 835 TTACGTGGCAGATTTGGCAGAGATCTCCCATGGGTTAAACCATGG 884
 101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyLys 117
 |||||
 885 CTGCAATTAACCAAGAAAGCCAGGTCCCGCAATTCGGCTTGAGAGAA 934
 117 sAsnLysHisSerAlaGluLysValGluProSerSerSerTyrLeuProA 134
 |||||
 935 AAACAAACACTCGGCTGAAAAAGTCAAGCCCTC.TCCTCATTCATACCC 983
 134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150
 |||||
 984 GAGATTGAGAGACCGCGCTTGCGCGAGAACCCCATCTGTTCCCCAC 1033
 151 ProLeuSerSerThrGlyCysCysGluGlyThr.SerAlaProProGlyA 167
 |||||
 1034 CCTTATTCGGGCACAGGTGCGCGAGGGGACCTTGGCCCTCGAGAG 1083
 167 laProValValGluGlyProAlaAlaGlyThrArgSerArgGlyAla 183
 |||||
 1084 CTCGGCGGGGAGGACCTGTCAGGAGCTCGGAGCGCGGAGGGCGCC 1133
 184 ThrProLysArgThrAspGluIleAlaIleLeuProLeuArgThrTyrG 200
 |||||
 1134 ACCCCGAGCGGACAGACAGATCGGACATTACCGCTGCGACAGTACGG 1183
 200 yProProMetLProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPhe 217
 |||||
 1184 CCGTCCACACCGCGGGGCCAATTCAGACCCCTCCAGTATGCGCCCTTT 1233
 217 eSerAlaAspLeuTyrAsnTrpLysThrAsnHisProProPheSerGlu 233
 |||||
 1234 CTCTGAGATCTCTATATTTGAAAACCTACCATCCCTTCTCGGAG 1283
 234 AspProGlnArgLeuThrGlyLeuValGluSerLeuMetPheSerHisG 250
 |||||
 1284 GATCCCAACAGCCCTCAGCGGGTGTGGAGTCCCTATGTTCTCACCAC 1333
 250 nProThrTrpAspAspCysGlnGlnLeuGlnTyrTrpPheThrThg 267
 |||||
 1334 GCCTACTTGGGATGATGATTCACACAGCTGTCAGACACTCTTCACAC 1383
 267 LuGluArgGluArgIleLeuLeuGlnAlaArgLysAsnValProGlyAla 283
 |||||
 1384 AGGAGCGAGAGATTCATTTAGAGCTTGAAAATAATGTTCTCGGGGCC 1433
 284 AspGlyArgProThrGlnLeuGlnAsnGluIleAspMetClyPheProLe 300
 |||||
 1434 GACGGGCGACCCACAGCGGTGTCAAAATGAGATTGACATGGGATTTCCCT 1483


```

300 uThArgProGlyTTrpAspTyrAsnThrAlaGluGlyArgGluSerLeuL 317
    |||||||
1484 AACTGCGCCCGGTGGAGCTACACACGCGCTGAAGGTAGGAGAGCTTGA 1533
    |||||||
317 ySlleTyrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArg 333
    |||||||
1534 AAATCTATCGCCAGGCTCTGTGGGCTCCGGGGCGCTCAAGACGG 1583
    |||||||
334 ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGluPr 350
    |||||||
1584 CCCACTATTTGGCTTAAGTAAAGAAAGATGATGACGAGACCGAATGAC 1633
    |||||||
350 oProSerValPheLeuGluArgLeuMetGluAlaPheArgArgPheThrP 367
    |||||||
1634 CCCCTCTGTTTCTTCTGAGAGGCTCTTGAAGCCTTCAGCGGGTACACC 1683
    |||||||
367 ropheAspProThrsSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 383
    |||||||
1684 CTTTGATCCCACTCAGAGGCCCAAAAGCCCTCAGTGGCTTGGCTTT 1733
    |||||||
384 lLeGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGlu 400
    |||||||
1734 ATAGGACAGTCAGCCTTGATATTAGAAAGACCTTCAGAGACTGGAAAG 1783
    |||||||
400 yLeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluLysValT 417
    |||||||
1784 GTTACAGAGAGCTGAGTTACGTGATCTAGTGAAGAGAGCAGCAAGTAT 1833
    |||||||
417 yTTrpArgArgGluThrGluGluGluLysGluGlnArgLysGluLysGlu 433
    |||||||
1834 ATTACAAAGGGAGACAGAAAGAAAGAAAGCAAAAGAAAGGAGACAA 1883
    |||||||
434 ArgGluGluArgGluGluArgArgAspArgArgGluGluLysAsnLeuTh 450
    |||||||
1884 AGAGAGAAAGGGAGAGAGAGCTATATAACGGCAAGACAGAAATTTGAC 1933
    |||||||
450 rLysIleLeuAlaIleValAlaGluGlyLysSerSerArgGluArgGlu 467
    |||||||
1934 TAAGATCTTGGCTGCAAGTGTGMAAGGAAAGCAATACGCAAGAGAGA 1983
    |||||||
467 rGAspPheArgLysIleArgSerGlyProArgIleSerLysAsnLeuGly 483
    |||||||
1984 GAATTTTGGCAAAATTAGGTCAAGGCCCTAGACAGTCAAGGAACCTGGGC 2033
    |||||||
484 AsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysG 500
    |||||||
2034 AATAGGACCCCACTCGACAAAGACCAATGTGCATATTGTAAGAAGAGAG 2083
    |||||||
500 yHISrTPAlaArgAsnCysProLysLysGlnLysAsnLysGlyProLysValL 517
    |||||||
2084 ACACTGCGGCAAGAACTGCCCAAGAGGAAACAAAGAGCAAGAGATCC 2133
    |||||||
517 euAlaLeuGluGluAspLysAsp 524
    |||||||
2134 TACCTTAGAAGAAAGATTAAAGAT 2156
    |||||||
seq_name: /SIDS2/gcdata/geneseq/geneseq/NA1997.DAT: AAT74811
seq_documentation_block:
ID AAT74811 standard; cDNA; 8060 BP.
XX
AC AAT74811;
XX
DT 11-FEB-1998 (first entry)
XX
DE Porcine retrovirus Tsukuba-1 cDNA.
XX
KW Retrovirus; porcine; GAG protein; POL protein; ENV protein;
KW xenotransplantation; infectious; provirus; organ transplant; donor;
KW activated virus; Tsukuba-1; PCR; ss.
XX
OS Porcine retrovirus.

```

```

XX Key Location/Qualifiers
FH 86..2002
FT CDS /*tag= a
FT /*note= "putative ENV protein"
FT 3112..4686
FT /*tag= b
FT /*note= "putative GAG protein"
FT 4871..8060
FT CDS /*tag= c
FT /*note= "putative POL protein (partial)"

WO921836-A1.
19-JUN-1997.
13-DEC-1996; 96WO-US19680.
14-DEC-1995; 95US-0572645.
(GENHO ) GEN HOSPITAL CORP.
Fishman JA:
WI: 1997-332804/30.

New nucleic acid from porcine retroviruses - used for detecting
PT viruses in transplant or other tissue and for assessing risk of
PT transmitting infection to graft recipient
XX
PS Claim 1; Fig 1; 128bp; English.
XX
CC This sequence represents the purified porcine retroviral cDNA
CC sequence of Tsukuba-1 and contains the putative coding regions for viral
CC proteins GAG, POL and ENV. This sequence and PCR fragments generated
CC from the sequence (see AAT74812-T74882) could be used to screen organs
CC for porcine retroviruses prior to xenotransplantation. Transplantation
CC can increase the likelihood of retroviral activation if intact and
CC infectious proviruses are present. The porcine retroviral sequence can
CC be used to generate probes to determine the level (e.g. copy number) of
CC intact (i.e. potentially replicating) porcine provirus sequences in a
CC strain of xenograft transplantation donors. It can be used to detect
CC mutations, genetic lesions or viral recombinants and to determine the
CC histological localization of activated retroviruses. Using polymerase chain
CC Reaction DNA Quantitation (PQ) on blood mononuclear cells, infectivity
CC titration and susceptibility testing can be performed. Ultimately animal
CC donors without intact porcine retroviral sequences or with a lower copy
CC number of viral elements could be selected.
XX
SQ Sequence 8060 BP; 2233 A; 1959 C; 2012 G; 1856 T; 0 other;

```

```

alignment_scores:
Quality: 2634.00 Length: 525
Ratio: 5.105 Gaps: 1
Percent Similarity: 98.286 Percent Identity: 94.286

```

alignment_block:

US-09-171-553B-4 x AAT74811 ..

Align seg 1/1 to: AAT74811 from: 1 to: 8060

```

1 MetGlyGlnPheValThrThrProLeuSerLeuThrLeuAspHisTyrPth 17
    |||||||
3112 ATGGGACAGACGGGTGACGACCCCTCTTAGTTGACTCTCGACCATTTGAC 3161
    |||||||
17 rGluValArgSerArgAlaHisAsnLeuSerValGlnValLysGlyP 34
    |||||||
3162 TGAAGTTAAATTCAGGCGCTCATATATTGTCAGTTGAGGTTAAAGAGGAC 3211
    |||||||
34 rortPpgGlnThrPheCysAlaSerGluThrProThrPheAspValGlyTyr 50
    |||||||
3212 CTGGCAGACTTCTGTCTGTGAATGCGCCGACATTTCGATTGGATGG 3261
    |||||||

```

51 ProSerGluGlyThrPheAsnSerGluIleLeuAlaValLysAlaIle 67
 |||||
 3262 CCATCAGAGGGGACCTTAAATTCAGATTAATCTGGCTGTAAAGCAAT 3311
 |||||
 67 eilePheGluThrGlyProGlySerHisProAspGluGluProTyrIle 84
 |||||
 3312 TATTTTTCAGCTGACCGGGCTCTCATCCGATCAGGAGCCCTATATCC 3361
 |||||
 84 eutHTrpGluAspLeuAlaGluAspProProTyrValLysProTyr 100
 |||||
 3362 TTACGTGGCAGATTGGCAGAGATCTCCGCATGGTTAAACCATGG 3411
 |||||
 101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGly 117
 |||||
 3412 CTGAATAGCCAGAAAGCCAGGTCGCCGAATTCGTGCTTGAGAGAA 3461
 |||||
 117 SASLysHisSerAlaGluLysValGluProSerSerSerTyrLeuPro 134
 |||||
 3462 AACCAACACTCGGCTGAAAAAGTCMAAGCCCTC.TCCTCATATCTACCC 3510
 |||||
 134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150
 |||||
 3511 GAGATTGAGAAACACCGGCTGGCCGGAACCCCAATCTGTTCCGCCACC 3560
 |||||
 151 ProLeuSerSerThrGlyCysCysGluGlyThr.SerAlaProProGly 167
 |||||
 3561 CCTTATCTGGACAGGGGTGCCGAGGGGACCCCTTGGCCCTCTGGAG 3610
 |||||
 167 lAProValAlaGluLysProAlaAlaGlyThrArgSerArgArgGlyAla 183
 |||||
 3611 CTCGGCGGTGGAGGACCTTCTGCAGGACTCGGAGCCGAGGGGCGCC 3660
 |||||
 184 ThrProGluArgThrAspGluIleAlaIleLeuProLeuArgGlyTrpGly 200
 |||||
 3661 ACCCGGAGCGGACAGAGATCGGACATTACCGCTGCGCAGCTACGG 3710
 |||||
 200 yProMetProGlyGlyGluLeuGluProLeuGlnTyrTrpProPhe 217
 |||||
 3711 CCTCCACACCGGGGGGCCAATTGCAAGCCCTCCAGATTTGGCCCTTT 3760
 |||||
 217 eSerAlaAspLeuTyrAsnTrpLysThrAsnHisProProPheSerGlu 233
 |||||
 3761 CTTCTGCAGATCTATATATGGAAACTAACCATCCCTCTTCGAGG 3810
 |||||
 234 AspProGluArgLeuThrGlyLeuValGluSerLeuMetPheSerHisGly 250
 |||||
 3811 GATCCCAACGCTCAGCGGGGTGGAGTCCCTTATGTTCTCTCACCA 3860
 |||||
 250 nProThrTrpAspAspCysGluGluLeuLeuGlnThrLeuPheThrThr 267
 |||||
 3861 GCCTCTTGAGATGATGATTCACACAGCTGTCGACACACTTCACAAACG 3910
 |||||
 267 lGluArgGluArgIleLeuLeuGluAlaArgLysAsnValProGlyAla 283
 |||||
 3911 AGGAGCGAAGAGAAATTCATTAGAGGCTAGAAAAAATGTTCTCGGGCC 3960
 |||||
 284 AspGlyArgProThrGluLeuGlnAsnGluLeuAspMetGlyPheProLe 300
 |||||
 3961 GACGGGCGACCCACGCGGTTGCCAAATGAGATTGACATGGATTTCCTT 4010
 |||||
 300 uThrArgProGlyTrpAspTyrAsnThrAlaGluGlyArgGluSerLeu 317
 |||||
 4011 AACTGCCCGCGGTGGAGTACACACGCTCAAGGTAGGAGACTTGA 4060
 |||||
 317 ySLeTyrArgGluAlaLeuValAlaGlyLeuArgGlyAlaSerArgArg 333
 |||||
 4061 AATCTATGCGCAGGCTCTGTGGCGGCTCCGGGCGCCCTCAAGACGG 4110
 |||||
 334 ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGluPr 350
 |||||
 4111 CCCACTAATTTGGCTAAGAGAGATGATGACGGGACCCGAAATGAAACC 4160

350 oProSerValPheLeuGluArgLeuMetGluAlaPheArgArgPheThr 367
 |||||
 4161 CCCCTGTGTTTCTTGAGAGGCTCTTGAGACCTTCAGGGCGGTACACC 4210
 |||||
 367 roPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 383
 |||||
 4211 CTTTGATCCCACTCAGAGGCCCAAAAAGCCTCAGTGCTTGGCTTT 4260
 |||||
 384 lIleGlyGlnSerAlaLeuAspIleArgLysLysLeuGluArgLeuGlu 400
 |||||
 4261 ATAGGACAGTCAAGCCTTGATATTAGAAAGAAAGCTTCAGAGCTGGAAG 4310
 |||||
 400 yLeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGlyValTr 417
 |||||
 4311 GTTACAGAGAGCGTGAATTCGTGATCTAGTGAAGAGAGCAGAAATAT 4360
 |||||
 417 yTrpArgArgGluTrpGluGluGluGlyGluAlaArgLysLysGlu 433
 |||||
 4361 ATTACAAAAGCGAGACAGAAAGAAAGGAAAGCAAAAGAAAGAGAGAA 4410
 |||||
 434 ArgGluGluArgGluGluArgArgAspArgArgGlnGluLysAsnLeuTh 450
 |||||
 4411 AGAGAGAAAGAGAGAAAGACCTATATAAGGCAAGAGAAATTTGAC 4460
 |||||
 450 rLysIleLeuAlaAlaValAlaGluGlyLysSerSerArgGluArgGlu 467
 |||||
 4461 TAAGATCTTGCTGCTGAGTGTGAAGGAAAGCAATACGAAAGAGAA 4510
 |||||
 467 rGAspPheArgGlyLysArgSerGlyProArgGlnSerGlyAsnLeuGly 483
 |||||
 4511 GAGATTTTGAAGAAATTAAGTTCAGGCCCTTAGACAGTCAGGAAACCTGGGC 4560
 |||||
 484 AsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGly 500
 |||||
 4561 AATAGAGACCCACTCGACAGAGACCAATGTCATATTGTAAGAAAGAGAG 4610
 |||||
 500 yHisTrpAlaArgAsnCysProLysLysGlyAsnLysGlyProLysVal 517
 |||||
 4611 ACACTGGGCAAGAAAGCTGCCCAAGAGGAAAGCAAAAGGACCAAGGATCC 4660
 |||||
 517 euAlaLeuGluGluAspLysAsp 524
 |||||
 4661 TAGCTCTAGAAAGATGAAGAT 4683
 |||||

seq_name: /SID2/gc9data/geneseq/geneseqn/NA2001.DAT:AAF77725
 seq_documentation_block:
 ID AAF77725 strand:..cDNA: 8060 BP.
 XX AAF77725;
 AC
 XX
 DT 23-MAY-2001 (first entry)
 XX
 DE Tsukuba-1 cDNA.
 XX
 KW Tsukuba-1; retrovirus; graft transplantation; xenotransplantation; ss.
 XX
 OS Porcine retrovirus.
 XX
 PN US6190861-B1.
 XX
 PD 20-FEB-2001.
 XX
 PF 13-DEC-1996; 96US-0766528.
 XX
 XX 14-DEC-1995; 95US-0572645.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PI Fishman JA;
 XX
 DR WPI; 2001-256211/26.

PT Assessing risk of endogenous retroviruses in clinical practice and in
PT xenotransplantation, comprises using probe sequences derived from swine
PT or miniature swine retroviral genome -
XX
PS Claim 1; Fig 1; 127pp; English.
XX
CC The present invention relates to a method for screening a cell or tissue
CC for the presence or expression of a retrovirus (RV), comprising
CC contacting a target nucleic acid from the cell or tissue with a second
CC nucleic acid from the present invention (e.g. the present sequence or a
CC fragment thereof). The method is useful for RV detection and to assess
CC graft transplantation risk. Screening of animals allows the elimination
CC of donors with active replication of known viruses. Inactive proviruses
CC can be detected and inactivated, allowing identification and elimination
CC of potential human pathogens derived from swine in a manner not possible
CC in the outbred human organ donor population and is important to the
CC development of human xenotransplantation.
XX
SQ Sequence 8060 BP; 2233 A; 1959 C; 2012 G; 1856 T; 0 other;

alignment_scores:
Quality: 2634.00 Length: 525
Ratio: 5.105 Gaps: 1
Percent similarity: 98.286 Percent identity: 94.286

alignment_block:
US-09-171-553B-4 x AAF77725 ..

Align seg 1/1 to: AAF77725 from: 1 to: 8060

1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisTrpTh 17
|||||
3112 ATGGGACGAGCGGTGACGACCCCTTGTAGTTCGACCTCGACCATTTGGAC 3161
17 TGIuValArgSerArgAlaHisAsnLeuSerValGlnValLysGlyP 34
|||||
3162 TGAAGTTAAATCCAGGCGCTCAATTTGTGACATTGAGTTAAGAAAGGAC 3211
34 roTrpGlnThrPheCysAlaSerGluTrpProThrPheAspValGlyTrp 50
|||||
3212 CTGGCAGACTTCTGTCTGTGATGAGCCGACATTCGATGATGGAGGG 3261
51 ProSerGlnGlyThrPheAsnSerGluLeuLeuAlaValLysAlaI 67
|||||
3262 CCATGAGAGGGGACCTTAATTCGAGATTAATCCGCGCTGTTAAGCAT 3311
67 eIlePheGlnThrGlyProGlySerHisProAspGlnGluProGlyLe 84
|||||
3312 TATTTTTCAGACTGACCCGGCTCTCATCCGATCAGAGGCCATATATCC 3361
84 euThrTrpGlnAspLeuAlaGluAspProProProThrPheValLysPro 100
|||||
3362 TTACGTGGCAGAGATTGGCAGAGATCTCCGCCATGGGTTAAACCATGG 3411
101 LeuAsnLysProArgLysProGlyProArgGluLeuAlaLeuGlyGly 117
|||||
3412 CTGAATTAAGCCAGAAAGCCAGGTCGCCGAATTCGGCTTTGGAGAAA 3461
117 sAsnLysHisSerAlaGlyLysValGluProSerSerSerTrpLeuPro 134
|||||
3462 AAACAAACACTCGCTGAAAAAGTCAGCCCTC; TCCTCATATCTACCCC 3510
134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150
|||||
3511 GAGATTGAGGAACACACCGCTTGGCGGAGAACCCCATCTGTTCCCCACC 3560
151 ProLeuSerSerThrGlyCysGlyGlyThr; SerAlaProProGlyAla 167
|||||
3561 CCTTATCTGCGACAGGCTGCCGAGGGGACCCCTTGGCCCTCTCTGAG 3610
167 laProValValGlnGlyProAlaAlaGlyThrArgSerArgArgGlyAla 183
|||||
3611 CTCGGCGGTGAGGAGACTTTCAGAGGACTCGGAGCCGGAGGGCCGCC 3660
184 ThrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTrpG 200
|||||
3661 ACCCGGAGGGGACAGAGATCGGACATTCACCGCTGGCGACGTACGG 3710
200 yProProMetProGlyGlyGlnLeuGlnProLeuGlnTrpTrpProPhe 217
|||||
3711 CCTCCACACCGGGGGCCCAATTCAGAGCCCTCCAGTATTCGCTTTT 3760
217 eSerAlaAspLeuTrpAsnTrpLysThrAsnHisProProPheSerGlu 233
|||||
3761 CTCTTCGAGATCTCTAATAATGGAAAACTAACCATCCCTTCCTGGGAG 3810
234 AspProGlnArgLeuThrGlyLeuValGluSerLeuMetPheSerHisG 250
|||||
3811 GATGCCAACAGCCTCAGGGGTTGGTGGAGTCCCTATGTTCTCTACCA 3860
250 nProThrTrpAspAspCysGlnGlnLeuGlnThrLeuPheThrTrpG 267
|||||
3861 GCTACTTGGGATGATGTCACACAGCTGTCAGACACTCTCACAACCG 3910
267 luGluArgGluArgIleLeuLeuGluAlaArgLysAsnValProGlyAla 283
|||||
3911 AGGAGCGAGAGAAATTCATTAAGAGCTAAGAAAAATGTTCTGGGGCC 3960
284 AspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPhePro 300
|||||
3961 GACGGGGGACCCACCGCGTTGCAAAATGAGATTGACATGGGATTTCCCT 4010
300 uThrArgProGlyTrpAspTrpAsnThrAlaGluGlyArgGluSerLeu 317
|||||
4011 AACTGCGCCCGTTGGGACTACAAACGCGCTGAAGTAGAGGAGAGCTTGA 4060
317 ySileTrpArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArg 333
|||||
4061 AAATCTATCCGACGCTGCTGGTGGCGGCTCTCGGGGCGCTCAACACGG 4110
334 ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGlu 350
|||||
4111 CCCACTATTTGGCTAAGTAAGAGAGATCATCAGAGACCGGATGAGAC 4160
350 oProSerValPheLeuGlnLysLeuMetGluAlaPheArgArgPheThr 367
|||||
4161 CCCCTGTTTCTTCTGAGAGGCTCTGGAAGCCTCAGCGGCTACACCC 4210
367 roPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 383
|||||
4211 CTTTGATCCCACTCAGAGGCCCAAAAGCCCTCAGTGGCTTTGGCTTT 4260
384 IleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGlu 400
|||||
4261 ATAGGACAGTACGCTTGATATTAGAAAGAACTTCAGAGACTTGGAGG 4310
400 yLeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluLysVal 417
|||||
4311 GTTACAGAGAGCTGAGTACGTATCTACGAGAGGCGCAGAAAGATAT 4360
417 yTrpArgArgGluThrGlnGluGlyGlnGluArgLysGluLysGlu 433
|||||
4361 ATTAACAAGAGGAGACAGAAAGAAAGGAGCAACAAAGAAAGAGAGAA 4410
434 ArgGluGluArgGluGluArgAspArgArgGlnGluLysAsnLeuTh 450
|||||
4411 AGAGAGAAAGGAGGAGAAAGACGTAATAACGGCAAGAGAAATTTGAC 4460
440 rLysIleLeuAlaAlaValValGluGlyLysSerSerArgGluArgGlu 467
|||||
4461 TAAGATCTTGGCTGCACTGTTAAGAGGAAACCAATACGAAAGAGAGA 4510
467 rGAspPheArgLysIleArgSerGlyProArgGlnSerGlyAsnLeuGly 483
|||||
4511 GAGATTTTAGGAAATTTAGGTCAAGCCCTAGACAGTACAGGAAACTGGC 4560

```

484 AsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysG1 500
|||||
4561 AATAGGACCCCACTCGACAGACCAATCTGCATATTGTAAGAAAGAG 4610
500 yHISTPAlaArgAsnCysProLysLysGlnLysGlyProLysValL 517
|||||
4611 ACACCTGGCAAGCACTGCCCAAGAGGAAACAAAGAGGAGGAGGATCC 4660
517 euAlaLeuGluGluAspLysAsp 524
|||||
4661 TAGCTCTAGAGAGATTAAGAT 4683

seq_name: /SIDS2/gcdata/geneseq/geneseq/NA1999.DAT:AAV69750
seq_documentation_block:
ID AAV69750 standard; cDNA; 8655 BP.
XX
XX AAV69750:
XX
XX 01-MAR-1999 (first entry)
XX
XX Mus dunni endogenous virus.
XX
XX MDEV; retrovirus; packaging cell line; gene transfer; gene therapy;
XX vector; ss.
XX
XX Mus dunni endogenous virus.
XX
XX Key location/Qualifiers
XX CDS 422..2290
XX /tag= a
XX /transl_except= (pos:1..2, aa:Met)
XX /product= Gag (glycosylated)
XX CDS 722..2290
XX /tag= b
XX /product= Gag
XX CDS 2291..5902
XX /tag= c
XX /product= Pol
XX CDS 5754..7775
XX /tag= d
XX /product= Env
XX
XX W09850538-A1.
XX
XX 12-NOV-1998.
XX
XX 08-MAY-1998; 98WO-US09452.
XX
XX 08-MAY-1998; 98US-0075272.
XX 09-MAY-1997; 97US-0046140.
XX
XX (HUTC-) HUTCHINSON CANCER RES CENT FRED.
XX
XX Bonham L, Miller AD, Wolgamot G;
XX
XX WPI: 1999-034718/03.
XX P-PSDB: AAM81570-73.
XX
XX New retroviral packaging cells - containing Mus dunni endogenous
XX virus sequences to target cells, retrovirus gag and pol genes and a
XX heterologous gene of interest.
XX
XX Claim 3; Page 60-64; 85pp; English.
XX
XX This is the complete nucleotide sequence of Mus dunni endogenous
XX virus (MDEV) including open reading frames encoding glycosylated
XX Gag (see W815700, Gag (see AAM81571), Pol (see AAM81572) and Env (see
XX AAM81573). A cultured packaging cell is claimed which produces a
XX replication-defective retroviral vector (RDV) particle, where
XX the packaging cell is a vertebrate cell capable of expressing and
XX assembling retroviral proteins, comprising: (a) a first vector

```

```

CC encoding a retroviral envelope protein having amino acid residues
CC MDEV that direct binding of the retroviral particle to MDEV
CC retroviral receptors on a target cell; and (b) a second vector
CC encoding retrovirus gag and pol proteins, where upon expression of
CC the vectors in the packaging cell in the presence of a vector
CC having a sequence of a heterologous gene of interest, a
CC replication-defective retroviral particle is produced that binds to
CC MDEV receptors of target cells. Also claimed are: (1) a cultured
CC packaging cell for producing a RDV particle; (2) methods for
CC producing a RDV particle comprising a heterologous gene of
CC interest; (3) cultured packaging cell line PD22; and (4) a RDV
CC produced by a method as in (2). The MDEV receptor is present on a
CC variety of cells rendering MDEV pseudotype packaging cells useful
CC in methods of mammalian and particularly human gene transfer for
CC gene therapy. The MDEV packaging cells are a stable and
CC reproducible source of retroviral particles. Clones may be
CC isolated from these populations that produce high titre virus. The
CC packaging cell lines may be selected and cloned for other desirable
CC properties, such as stability of in vivo growth, lack of production
CC of helper virus, lack of reinfection by viral particles packaged in
CC the cell, stability from genetic rearrangement and recombinational
CC events, resistance to complement lysis, and improved ability to
CC infect cells from higher mammals.
XX
XX S0 Sequence 8655 BP; 2316 A; 2228 C; 2051 G; 2060 T; 0 other:

```

```

alignment_scores:
Quality: 1684.00 Length: 541
Ratio: 3.907 Gaps: 10
Percent Similarity: 79.667 Percent Identity: 61.553

```

```

alignment_block:
US-09-171-553B-4 x AAV69750 ..

```

```

Align seg 1/1 to: AAV69750 from: 1 to: 8655

```

```

1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisTrpTh 17
|||||
722 ATGGGACAAATCTGTCTCCACGCCCCCTTCTTAACCTGGAGCATTTGGAA 771
17 G1uValArgSerArgAlaHisAsnLeuSerValGlnValLysGlyP 34
|||||
772 GGAGGTAAATAATCAGACACACATCATCATGAGGAGTTAGGAAAGGCC 821
34 rOTrPGInThrPheCysAlaSerGlnTrpProThrPheAspValGlyTrp 50
|||||
822 CATGGCAAACTTTTGGCGCTCCGAGTGGCCAGTTTGGAGTGGGCTGG 871
51 ProSerGluGlyThrPheAsnSerGluIleIleLeuAlaValLysAlaI 67
|||||
872 CCACCGGAGGGTGTCTTTGACTTGTCTCATGATCGCCGCCCTCAGGCAAT 921
67 e1lePheG1InThrClyProGlySerHisProAspGlnGluProTyrIle 84
|||||
922 TGTTTTTCAGGAA...GAGGGGGGTCACTCATCATCATCTCCATCATTTG 968
84 eutHrTrPGInAspLeuAlaGluAspProProThrProValLysProTrp 100
|||||
969 TGACCTGGCAGAGTCTGTGTCAGTCCACCTTCGTGGGTCAAGACCTGG 1018
101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyLys 117
|||||
1019 .....ACCCCAATTCCTTCGAATCTAGCGGTTCACAGTTGC 1053
117 sAsnLysHisSerAlaGluLysValGluProSerSer.....SerTyrL 132
|||||
1054 CCAATTCGAGAGCTGAGAAATCTGTGTCGTCAGACCCCAAGATTTT 1103
132 eutProArgAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSer 148
|||||
1104 ATCCAGAGATTGACGACCTCTCTGTGATGATGCCAACCTCCCTTAC 1153

```


CC retroviruses with cell membranes, while maintaining mutant envelope
 CC protein incorporation into a virion, and viral titers of about two
 CC orders of magnitude within that observed for wild-type retrovirus when
 CC the protein or fragment is expressed on the surface of a retroviral
 CC particle. The proteins have an increased ability to penetrate targets,
 CC typically cells and a correspondingly increased ability to deliver
 CC nucleic acids or drugs. The mutated nucleic acid is useful for gene and
 CC drug therapy, especially as drug delivery vehicles. The retrovirus
 CC particles can be utilized to transduce eukaryotic cells. The transduced
 CC cells are useful in the treatment of cancer in a human. Other diseases
 CC contemplated for treatment include adenosine deaminase deficiency (ADA),
 CC thalassemia, hemophilia, diabetes, alpha anti trypsin deficiency, brain
 CC and neural disorders, phenylketonuria, growth disorders, heart diseases
 CC and immune diseases.

SO Sequence 8088 BP; 2103 A; 2251 C; 1982 G; 1752 T; 0 other;

alignment_scores:

Quality: 1660.50 Length: 542
 Ratio: 3.944 Gaps: 10
 Percent Similarity: 77.675 Percent Identity: 61.439

alignment_block:

US-09-171-553B-4 x AA245540 ..

Align seg 1/1 to: AA245540 from: 1 to: 8088

1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisThrPth 17
 631 ATGGGACAAAGATTAATCTACCCCTATCTCCCTACCTAAATCACTGGAG 680
 17 rGluValArgSerArgAlaHisLeuSerValGlnValLysLysGlyP 34
 681 AGATGGAGAACAGAGGCTCAATCTATCCCTGGAATCAAAAGGAA 730
 34 rGTPGlnThrPheCysAlaSerGluTTPProThrPheAspValGlyTTP 50
 731 AATGGCAGACTTCTGTTCTCCGAGTGGCCACATTCGGCGGGGTGG 780
 51 ProSerGluGlyThrPheAsnSerGluLeuLeuAlaValLysAla11 67
 781 CCACGGAGGAGACTTTTAATCTCTGTCATTTTGGCATTAAGAT 830
 67 eilePheGlnThrGlyProGlySerHisProAspGlnGluProTyrLeu 84
 831 TGTCTTTCAGGAGAAC...CGCGACATCCGAGCACAGTCCATATACG 877
 84 eutThrTPGlnAspLeuAlaGluAspProProTTPProTTPValLysProTTP 100
 878 TGGTATGGCAGGACCTCGCCAGAAATCCGCCACATGGGTG..... 918
 101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyLys 117
 919CCAGCCTCCGCAAGGTCGCTGTGTCTGATAC 953
 117 sAsnLysHisSerAlaGluLysValGluProSerSerSerTyrLeuProA 134
 954 CCGAAGACAGATTGCGGGGAGG.....CCATCAGCTCCTCCCGACCC 997
 134 rAspArgGlyAlaAsp.....LeuAlaGlyThrProThr... 146
 998 CCATCTACCCGCAACAGACTTACTCTCTCTGAAACCCACAGCC 1047
 147CysSerProThrProLeuSerSerThrGlyCysCys 158
 1048 CCGCCCTATCCGCGGCACTGCCACCCCTTGCGCCCTCAGAGCATGG 1097
 158 sGlnGlyThrSerAlaProGlyAlaProValAlaGlnGlyProAla 175
 1098 ACCGCGCTCAGGCGAGATGCGGATATGATGATCCTGAGGGGCGACG 1147
 175 lAgGlyThrArgSerArgArgGlyAlaThrPro.....GluArgThr 188

1148 CTGGAGCAGAGAGTCGCGCCGCGAGTCACAGCAACTCGGCTCT 1197
 189 AspGluLeuAlaIleLeuProLeuArgThrTyrGlyProProMetProG 205
 1198 GACTCCACTGTGATTTTGGCCCTCCGAGCCATAGAGCCCGGCGAGCC 1247
 205 yGlyGlnLeuGlnProLeuGlnThrTTPProPheSerSerAlaAspLeu 222
 1248 CAATGGCCTGTGCTTACAAATATGCGCTTTTCTCCAGAGATCTT 1297
 222 yAsnThrPlyThrAsnHisProProPheSerGluAspProGlnAlaGlu 238
 1298 ATAATGTGAATCTTAATCATCCCTTTTCTAAAAACCCAGAGGCTTC 1347
 239 ThrGlyLeuValGluSerLeuMetPheSerHisGlnProThrTPAspAs 255
 1348 ACGGGGCTCTTGAGTCTCTTATGTTCTCCATCAGGCCACTTGGAGCA 1397
 255 pCysGlnGlnLeuLeuGlnThrLeuPheThrThrGluArgGluArg 272
 1398 TTGGCAACAGTCTCTACAGATTTCTTTCACCACTAGAGCAAGGAAAGAA 1447
 272 lLeuLeuGlnAlaArgLysAsnValProGlyAlaAspGlyArgProThr 288
 1448 TTCCTCGAGGCGCCGCAAAATGCTCTTGAGACAAATGGGGCCCTACA 1497
 289 GlnLeuGlnAsnGluLeuAspMetGlyPheProLeuThrArgProGlyT 305
 1498 CAGCTCCAGAACCTCATTAATGAGGCTTCCCTCATAGACTCTACTG 1547
 305 pAspTyrAsnThrAlaGluArgGluSerLeuLysIleTyrArgGln 322
 1548 GGATTAACAACACAGCCAGGTAGAGGAGCTCTTCTGTTACCGCGCGGA 1597
 322 lAlaValAlaGlyLeuArgGlyAlaSerArgArgProThrAsnLeuAla 338
 1598 CTCTAGGCGAGCTCAAAAGGGGAGCTGGGCTCTTACCAATTTGGCT 1647
 339 LysValArgGluValMetGlnGlyProAsnGluProProSerValPhe 355
 1648 AAGTAAAGAGAGGCTCTGCAAGGAGCGGACAGAACCCCTCGGTTTCT 1697
 355 uGluArgLeuMetGlnAlaPheArgArgPheThrProPheAspProThs 372
 1698 AGAAGCGCTGATGGAGGCTTAAGAGATACACTCCCTTTGATCCCTCT 1747
 372 eArgGluAlaGlnLysAlaSerValAlaLeuAlaPheIleGlyGlnSerAla 388
 1748 CTGAGGAGACACAGGCTGCGGTGCGCATGCGCTTTATCGACAGTCAGCC 1797
 389 LeuAspIleArgLysLysLeuGlnArgLeuGlnGlyLeuGlnAlaG 405
 1798 CCAGATATCAAGAAAGAAATACAGAGGCTAGAGGGGCTCCAGAGCTATTC 1847
 405 uLeuArgAspLeuValArgGluAlaGluLysValTyrTyrArgGluT 422
 1848 CTTCAAGATTTTACTAAAGAGGCGAAGAAAGGTGTACCAATAAGACAGA 1897
 422 hrgGluGlnGluLysGluGlnArgLysGluLysGluArgGluArgGlu 438
 1898 CAGAGAGAGAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1947
 439 GluArgArgAspArgArgGlnGluLysAsnLeuThrLysIleLeuAla 455
 1948 AGCGGGCGCGATAGCCCAAGAAAGAAACCTTACTAAATTCGCGCGC 1997
 455 aValValGluGlyLysSerSerArgGluArgGluArgAspPheArgLys 472
 1998 AGTAGTA.....AGTAGAAG..... 2013
 472 lLeuArgSerGlyProArgGlnSerGlyAsnLeuGlyAsn..... 484

```

2014 ..GGGTCCACAGTAGGACAGACAGGGAACCTGAGCACACAGGCAAGAG 2061
485 .....ArgThrProLeuAspLysAspGlnCysAlaIleTyrCys 496
2062 AACACCTAGGATGAGAAAGACCTCCTACTAGACAAAGACAGTCGCTACTG 2111
496 sLysGlnLysGlnIleThrPalaIleArgAsnCysProLysLysGlnAsn...L 512
2112 TAAAGAGAAAGGGCCATTGGCAAGAGATGTCCCGAATAAACACCTCA 2161
512 ysgLyProLysValLeuAlaLeuGlu 520
2162 GAGAGCCAGCAAGTCTTACCCCTAGAT 2187

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT:AA073731
seq_documentation_block:
ID   AA073731 standard; DNA; 8535 BP.
XX
XX   AA073731:
AC   AA073731:
XX
XX   10-MAY-1995 (first entry)
XX
XX   GALV SEATO genome.
XX
XX   GALV; gibbon-ape-leukemia virus; retro virus; vector; plasmid p558;
XX   plasmid p521; plasmid p537; transgenic animal; gene therapy; ss.
XX
XX   Gibbon leukemia virus SEATO.
XX
XX   WO9423048-A.
XX
XX   13-OCT-1994.
XX
XX   06-APR-1994; 94MO-US03784.
XX
XX   06-APR-1993; 93US-0043311.
XX
XX   (USSH ) US SEC DEPT HEALTH.
XX   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX   Deacon NJ, Elden MV, Hooker DJ, Wilson CA.
XX
XX   WPI; 1994-333206/41.
XX
XX   Recombinant DNA contg. defective gibbon ape leukemia virus
XX   genome - and specific gene, also derived mammalian cells, virions
XX   etc., useful in gene therapy of e.g. cancer, viral infections or
XX   inherited diseases
XX
XX   Disclosure; Page 30-34; 68pp; English.
XX
XX   Novel replication-deficient retro virus vectors p558 (AA073733), p521
XX   (AA073734) and p537 (AA073735) include a GalV-derived packaging site,
XX   esp. the 200-910 base region of the GalV SEATO genome, and
XX   regulatory sequences, partic. the GalV 3' long terminal repeat
XX   (AA073732). The vectors allow transfer of genes for gene therapy and
XX   transgenic animal breeding.
XX
XX   Sequence 8535 BP; 2210 A; 2376 C; 2084 G; 1865 T; 0 other:

alignment_scores:
    Quality: 1622.50      Length: 538
    Ratio: 3.836          Gaps: 11
    Percent Similarity: 78.625    Percent Identity: 60.595

alignment_block:
US-09-171-553B-4 x AA073731 ..
Align seg 1/1 to: AA073731 from: 1 to: 8535
1 MetGlyInhrValThrThrProLeuSerLeuThrLeuAspHisTrpTh 17

```

```

1001 ATGGGACAAAGATATTTCTACCCCTATCTCCCTACTCTTAATATCATCTGAG 1050
17 rGluValArgSerArgAlaHisAsnLeuSerValGlnValLysLysGlyP 34
1051 AGATGTGAGAAACAAGGGCTCACAATCTATCCGTGAAATCAAAAAGGAA 1100
34 roTrpGlnThrPheCysAlaSerGluThrProThrPheAspValGlyTyr 50
1101 AATGGACAGCTTCTGTCTCTCCGAGTGGCCACATTCGGCGTGGGTGG 1150
51 ProSerGlnGlyThrPheAsnSerGluIleIleLeuAlaValLysAlaI 67
1151 CCAACGGGAGGAGACTTTTATCTCTCTGTCATTTTGGCACTTAAAAAGAT 1200
67 eIlePheGlnThrGlyProGlySerHisProAspGlnGluProTyrIle 84
1201 TGCTTTACAGAGAGAAC...GGGGACATCCGAGCAAGTTCATATATATG 1247
84 eutThrTrpGlnAspLeuAlaGluAspProProProThrValLysProTrp 100
1248 TGGTATGGCAGAGACCTCGCCAGAAATCCCAACATGGTG..... 1288
101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyLys 117
1289 .....CCAGCTCCGGCAAGGTCGCTGTCTGTCTGATAC 1323
117 sAsnLysHisSerAlaGlnLysValGluProSerSerSerTyrLeuPro 134
1324 CCGAGACCAAGTTCGGGGAGG.....CCATAGCTCTCCCGCACCC 1367
134 rgaSparGlyAlaAlaAsp.....LeuAlaGlyThrProThr... 146
1368 CCATCTACCCGACACAGAGACTTACTCTCTCTCTCAACCCACAGCC 1417
147 .....CysSerProThrProLeuSerSerThrGlyCysCys 158
1418 CCGCCCTATCCGGGCGACACTGCCACCCCTCGCCCTCAGGCGATCG 1467
158 sGlnGlyThrSerAlaProProGlyAlaProValValGlnGlyProAla 175
1468 ACCGCCGTACGCCAGATGCCGATAGTACGATCTCGAGGGGCCAGCC 1517
175 laGlyThrArgSerArgArgGlyAlaThrPro.....GluArgThr 188
1518 CTGGGACACAGAGATCCCGCGGCCGAGTCACAGACACACTCGGTCCT 1567
189 AspGlnIleAlaIleLeuProLeuArgThrTyrGlyProProMetPro 205
1568 GACTCCACTGTGATTTTGGCCCTCCGAGCCATAGACCCCGCGAGCC 1617
205 yGlyGlnLeuGlnProLeuGlnIleTyrThrProPheSerSerAlaAspLeu 222
1618 CAATGGCTGTGTCCTCTCAATATATGAGCTTTTCTCTCAGAGATCTTT 1667
222 yTrsAsnTrpLysThrAsnHisProProPheSerGlnAspProGlnArgLeu 238
1668 ATATATTGGAATCTATCATCCCTCTTTTCTGAAAACACAGAGGTCTC 1717
239 ThrGlyLeuValGlnSerLeuMetPheSerHisGlnProThrTrpAsp 255
1718 ACGGGGCTCTTGTAGTCTCTATGTTCTCCCATACAGCCACTTGGAGCA 1767
255 pcysGlnGlnLeuLeuGlnIleThrLeuPheThrThrGlnGluArgGluArg 272
1768 TTGCCACACAGCTCTTACAGATTCTTTTCCACCTGAGGAACGGAGAA 1817
272 leuLeuGlnLualaArgLysAsnValProGlyAlaAspLysArgProThr 288
1818 TTCTCTCGAGGCGCCGCAAAATGTCTTCTTGGGAGCAATGGGCGCCCTACA 1867
289 GlnLeuGlnAsnGlnIleAspMetGlyPheProLeuThrArgProGlyTyr 305

```


1868 CAGCTCGAAGACTCATTAATGAGCCTCCCTCAATGACCTCACTG 1917
305 PASPTyrAsnThrAlaGluArgGluSerLeuLysIleTyrArgGlnA 322
1918 GGATTACACACAGCCGAGGTAGGAGCCTCTGTGCTACCGCCGGA 1967
322 IAlaValAlaGlyLeuArgGlyAlaSerArgProThrAsnLeuAla 338
1968 CTCTAGTGGCAGCTCTCAAGGGGACCTGCGCTCTACCAATTGGCT 2017
339 LysValArgGluValMetGlnGlyProAsnGluProProSerValPheLe 355
2018 AAGGTAAAGAGAGCTCTTGACAGGACCGCAGAACCCCTTGCTTCTT 2067
355 uGluArgLeuMetGluAlaPheArgArgPheThrProPheAspProThrS 372
2068 AGAAGCGCTGATGGAGGCTATGAGAGATACACTCCGTTGATCCCTCT 2117
372 eArgAlaGlnLysAlaSerValAlaLeuAlaPheIleGlyGlnSerAla 388
2118 CTGAGGGGACACAGCAGCTGCGCTCCCATGGCTTATCGGACAGTCA 2167
389 LeuAspIleArgGlyLysLeuGlnArgLeuGlnGlyLeuGlnGlnAla 405
2168 CCAGATATCAAGAAAGTTACAGAGGCTAGAGGGCTCCAGGACTATTC 2217
405 uLeuArgAspLeuValArgGluAlaGlnLysValTyrTyrArgArgGlu 422
2218 CTTACAAAGATTAGTAAAGAGGCGAAGAGGTGTACCATRAGAGAGAGA 2267
422 hArgGluGlnGluLysGlnGlnArgGlyLysGlnArgGluArgGlu 438
2268 CACAAGAGAGAAACACAAAGAGAAAGAGAGAGAGAGAGAGAGAGAG 2317
439 GluArgArgAspArgArgGlnGlnLysAsnLeuThrLysIleLeuAla 455
2318 AGGCGGCGCATAGCGCCGAGAGAAAGAACTTGACTAAATTTGCGCGC 2367
455 aValVal.....GluGlyLysSerSerArgGluArgGluArgAspPheA 470
2368 AGTAGTAGTAGTAGAGAGGCTCCACAGGTAGGACCCCTCCCGGAA.... 2413
470 rGlySIIeArgSerGlyProArgGlnSerGlyAsnLeuGlyAsnArgThr 486
2414CCTAGGGTAACTGACTGTG...GAGGGGAGCC 2443
487 ProLeuAsp.....LysAspGlnCysAlaTyrCysLysGlnLysGln 500
2444 CCCTATGTAGTCTCGTCAAGACCAAGTCCCATCTACTGTAAAGAGAGG 2493
500 yHisTrpAlaArgAsnCysProLysLysGlyAsn...LysGlyProLysV 516
2494 CCATTGGGCGAGAGATGTCCCGAAAAAACACGTCAAGAGAACCAAGG 2543
516 aLeuAlaLeuGln 520
2544 TTCTAGCCCTAGAT 2557

OM of: US-09-171-553B-5 to: N_Geneseq_1101:* out_format : pfs

Date: Feb 24, 2002 10:19 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Comugen Ltd.

```

# Fine parameters:
-MODEL=frame+pn.model -DEV=xlh
-Q/cgcn2.1/USPPO.spool/USO9171553/runat_22022002.122826_11203/app_query.fasta_1.3312
-DB=NGenesec_1101 -Qfmt=fastap -SUFFIX=feb23std.rng
-GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000
-LOOPEXT=0.000 -QGAPOP=4.500 -OCAPEXT=0.050 -XGAPOP=10.000
-XGAPEXT=0.500 -DEFAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000
-YGAPEXT=0.500 -DEFAPEXT=6.000 -DELEXT=7.000 -START=1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DALIGN=200
-THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pfs -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USR=USO9171553 @CGN1_1526 -NCP=6 -ICPU=3 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -NO_XLXPY -WAIT -THREADS=1

```

Search information block:

```
Database: N_Geneseq_1101: *
Database sequences: 930621
Database length: 428662619
Search time (sec): 338.2000
```

Sequence	Strd Orig	zScore	EScore Len	! Documentation
SI5D22/cgcdata/geneseq/NA1997	DAT:AAV09700	0.0	6304.00	8311.45
SI5D22/cgcdata/geneseq/NA1997	DAT:AAV09699	0.0	6284.00	8284.99
SI5D22/cgcdata/geneseq/NA1997	DAT:AAV74884	0.0	6099.50	8041.13
SI5D22/cgcdata/geneseq/NA2001	DAT:AAF77727	0.0	6099.50	8040.82
SI5D22/cgcdata/geneseq/NA2001	DAT:AAV67020	0.0	6095.50	8041.83
SI5D22/cgcdata/geneseq/NA2001	DAT:AAV67019	0.0	6095.50	8036.55
SI5D22/cgcdata/geneseq/NA2001	DAT:AAV67023	0.0	5858.00	7721.44
SI5D22/cgcdata/geneseq/NA2001	DAT:AAV67032	0.0	5646.50	7434.11
SI5D22/cgcdata/geneseq/NA1997	DAT:AAV74811	0.0	5637.50	7429.29
SI5D22/cgcdata/geneseq/NA2001	DAT:AAF77725	0.0	5637.50	7429.29
SI5D22/cgcdata/geneseq/NA2001	DAT:AAV67021	0.0	5635.50	7429.54
SI5D22/cgcdata/geneseq/NA1997	DAT:AAV09698	0.0	4840.00	6382.62
SI5D22/cgcdata/geneseq/NA2001	DAT:AAF77726	0.0	4489.50	5910.46
SI5D22/cgcdata/geneseq/NA1997	DAT:AAV74883	0.0	4489.50	5910.38
SI5D22/cgcdata/geneseq/NA1999	DAT:AAV69750	0.0	4339.50	5710.18
SI5D22/cgcdata/geneseq/NA2000	DAT:AAV65540	0.0	4319.00	5683.74
SI5D22/cgcdata/geneseq/NA1994	DAT:AAQ73731	0.0	4319.00	5683.18
SI5D22/cgcdata/geneseq/NA1996	DAT:AAQ73280	0.0	4000.50	5261.94
SI5D22/cgcdata/geneseq/NA1995	DAT:AAQ94266	0.0	3988.00	5214.56
SI5D22/cgcdata/geneseq/NA1997	DAT:AAQ81042	0.0	3980.00	5234.65
SI5D22/cgcdata/geneseq/NA1997	DAT:AAZ07063	0.0	3980.00	5234.65
SI5D22/cgcdata/geneseq/NA1997	DAT:AAZ08057	0.0	3980.00	5232.40
SI5D22/cgcdata/geneseq/NA1999	DAT:AAZ07067	0.0	3980.00	5232.40
SI5D22/cgcdata/geneseq/NA2001	DAT:AAQ89138	0.0	3979.50	5234.05
SI5D22/cgcdata/geneseq/NA1999	DAT:AAV7178	0.0	3970.00	5214.41
SI5D22/cgcdata/geneseq/NA2000	DAT:AAAR8061	0.0	3967.00	5215.94
SI5D22/cgcdata/geneseq/NA1997	DAT:AAAT9093	0.0	3967.00	5218.78
SI5D22/cgcdata/geneseq/NA1997	DAT:AAQ90694	0.0	3967.00	5218.78
SI5D22/cgcdata/geneseq/NA1997	DAT:AAQ90688	0.0	3967.00	5218.36
SI5D22/cgcdata/geneseq/NA1995	DAT:AAQ08452	0.0	3967.00	5217.43
SI5D22/cgcdata/geneseq/NA1996	DAT:AAQ07374	0.0	3967.00	5217.43
SI5D22/cgcdata/geneseq/NA1997	DAT:AAQ05061	0.0	3967.00	5217.43
SI5D22/cgcdata/geneseq/NA1998	DAT:AAV42715	0.0	3967.00	5217.43
SI5D22/cgcdata/geneseq/NA1998	DAT:AAV19511	0.0	3967.00	5217.43
SI5D22/cgcdata/geneseq/NA1998	DAT:AAV15346	0.0	3967.00	5217.43
SI5D22/cgcdata/geneseq/NA1998	DAT:AAV28143	0.0	3967.00	5217.43
SI5D22/cgcdata/geneseq/NA1999	DAT:AAAX40297	0.0	3967.00	5217.43
SI5D22/cgcdata/geneseq/NA2000	DAT:AAAZ00664	0.0	3967.00	5217.43
SI5D22/cgcdata/geneseq/NA2000	DAT:AAZ45539	0.0	3967.00	5217.43
SI5D22/cgcdata/geneseq/NA2000	DAT:AAZ45539	0.0	3967.00	5217.43

284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
 |||||
 3013 CACTTTTGGCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCGGCAG 3062
 |||||
 301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317
 |||||
 3063 CTCACCTGGACCGACTGCCCCAAGGGTTCAAGAACCTCCCGACCATCTT 3112
 |||||
 317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
 |||||
 3113 TGAGGAAGCCCTACACAGGAGCCTGGCCAACTTCAGAGATCCAAACCCCTC 3162
 |||||
 334 lnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350
 |||||
 3163 AGGTGACCTCCCTCCAGTACGTGGATGACCTGCTTCTGGGGGAGCCACC 3212
 |||||
 351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuGluLeuSerAs 367
 |||||
 3213 AAACAGGACTGCTTAGAAGGTACGAAGGCACCTACTGCTGGAATTTGCTGA 3262
 |||||
 367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
 |||||
 3263 CCTAGGCTACAGAGCCTCTGCTAAGAGGCCAGATTTGCAGAGAGAGG 3312
 |||||
 384 alThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400
 |||||
 3313 TAACATACTTGGGTACAGTTTGGGGGCGGCAGCGATGGCTGACGGAG 3362
 |||||
 401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG1 417
 |||||
 3363 GCACGGGAAGAAACTGTAGTCCAGATPACCGGGCCCCAACACACAGCCAAACA 3412
 |||||
 417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
 |||||
 3413 AGTCAGAGAGTTTTGGGCACAGCTGGATTTTGCAGACTGTGGATCCCGG 3462
 |||||
 434 lypPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450
 |||||
 3463 GGTTCGACCTTAGCAGCCCCACCTCTACCCGCTAACCAAGAAAGG 3512
 |||||
 451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467
 |||||
 3513 GGATTCCTCTGGGCTCTCAGCACACCAGAGGCAATTTGATGCTATCAAAA 3562
 |||||
 467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484
 |||||
 3563 GGCCCTGCTGAGCGCACCTGCTCTGGCCCTCCCTGACGTAACTAACCCCT 3612
 |||||
 484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500
 |||||
 3613 TTACCCCTTTATGCTGGATGAGCGTAAGGGAGTAGCCCGAGGAGTTTAAACC 3662
 |||||
 501 GlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLe 517
 |||||
 3663 CAAACCCCTAGGACCATGGAGAGACCTGTGTGCTACCTGCTCAAGAAGCT 3712
 |||||
 517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaLav 534
 |||||
 3713 TGATCTGTAGCCAGTGGTTGGCCGTATGCTGAAGGCTATCGCAGCTG 3762
 |||||
 534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550
 |||||
 3763 TGGCCATACTGGTCAAGGACGCTGACAAATTCGACTTTGGGACAGAAATA 3812
 |||||
 551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567
 |||||
 3813 ACTGTAATAGCCCCCATGCAPTGGAGAACATCGTTGCGCAGCCGCCAGA 3862
 |||||
 567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeu 584
 |||||
 3863 CCGATGGATGACCAACGCCCGCATGACCCACTATCAAGCCCTGCTCTCA 3912
 |||||
 584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600
 |||||

3913 CAGAGAGGGTCACTTTGCTCCACAGCGCTCTCAACCCCTGCCACTCTT 3962
 |||||
 601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617
 |||||
 3963 CTGCGCTGAAGAGACTGATGAACCACTGACTCATGATTCATCAACTATT 4012
 |||||
 617 uileGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634
 |||||
 4013 GATTGAGGAGACTGGGGTCCGCAAGGACCTTACAGACATACCGCTGACTG 4062
 |||||
 634 LyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650
 |||||
 4063 GAGAAGTGCTTAACCTGGTTCACTGACGGAAGCAGCTATGTGGTGAAGGT 4112
 |||||
 651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667
 |||||
 4113 AAGAGATGGCTGGGGCGGCAGTGGTGCACGGGACCCGACGATCTGGGC 4162
 |||||
 667 aserSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684
 |||||
 4163 CAGCAGCCTGCCGGAAGAACTTACGCCCAAAAGGCTGAGCTCATGGCCC 4212
 |||||
 684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700
 |||||
 4213 TCACCGCAAGCTTTGGGGCTGGCCGGAAGGAATCCATAAACATTTATACG 4262
 |||||
 701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717
 |||||
 4263 CACAGCAGGTATGCTTTTGGGACTGCACACGTACACGGGGCCATCTATAA 4312
 |||||
 717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734
 |||||
 4313 ACAAGGGGGTGGTTTACCTCAGCAGGAGGGAATAAAGAACAAAGAGG 4362
 |||||
 734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750
 |||||
 4363 AAATTCTAAGCCTATTAGAAGCCTTACATTTGCCAAAAGGCTAGCTATT 4412
 |||||
 751 IleHisCysProGlyHisGlnLysAlaLysAspLeuLeSerArgGlyAs 767
 |||||
 4413 ATACACTGTCTGGACATCAGAAAGCCAAAGATCTCATATCTAGAGGAA 4462
 |||||
 767 nGlnMetAlaAspArgValAlaLysGlnAlaGlnAlaValAsnLeuL 784
 |||||
 4463 CCAGATGGCTGACCGGGTTGCCAAGCAGCGCAGCCAGGCTGTTAACCTTC 4512
 |||||
 784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800
 |||||
 4513 TGCCTATAATAGAAACGCCCAAGCCCAAGAACCCAGACGACGACAGTACC 4562
 |||||
 801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817
 |||||
 4563 CTAGAAGACTGGCAAGAGATAAAAAGATAGACCAGTTCTCTGAGACTCC 4612
 |||||
 817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834
 |||||
 4613 GGAGGGGACCTGCTATACCTCATATGGGAAGGAAATCCTGCCCCACCAAG 4662
 |||||
 834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850
 |||||
 4663 AAGGTTAGATAATGTCCACAGATACATCGCTAACCCACCTAGGAAGT 4712
 |||||
 851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867
 |||||
 4713 AACACCTGCAGCAGTTGGTGGTGAAGCAATCCCTTATCATGTCTGAGGCT 4762
 |||||
 867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884
 |||||
 4763 ACCAGGAGTGGCTGACTCGGTGGTGAAGCAATTTGTGTCCTGCCAGCTGG 4812
 |||||
 884 alAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900
 |||||

4813 TTAATGCTAATCTTCCAGAAATACCTCCAGGAAGAGACTAAGGGGAGC 4862
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917
|||||
4863 CACCAGCGCTCACTGGAAGTGGACTTCACTGAGGTAAAGCGGCTAA 4912
917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934
|||||
4913 ATACGGAACAATAATCATCTATTGGTGTGTTAGACACCTTTTCAGGATGG 4962
934 aGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950
|||||
4963 TAGAGGCTTATCTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAGAA 5012
951 lLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967
|||||
5013 ATACTGGAGGAAATTTTCCAAAGATTGGAATACCTAAGGTAATAGGTC 5062
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984
|||||
5063 AGACAATGGTCCAGCTTTCGTGGCCAGGTAAGTCAGGAGCTGGCCAA 5112
984 lLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000
|||||
5113 TATTGGGATTCATTCGAAACTGCATTGTGCATACAGACCCCAAGCTCA 5162
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017
|||||
5163 GCACAGGTAGAGAGATCAATAGAACCATTAAGAGAGACCCCTTACCAAT 5212
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034
|||||
5213 GACCACAGAGACTGGCATTAATGATGGATGGCTCTCCCTGGCCCTTGTGC 5262
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050
|||||
5263 TTTTATGGGTGAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATGAA 5312
1051 LeuLeuTyrGlyGlyProProLeuAlaGluIleAlaPheAlaHisSe 1067
|||||
5313 TTGCTCTACGGGGGACCCCGCTTGGCAGAAATTTGCCCTTGCACATAG 5362
1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084
|||||
5363 TGCTGATGTGCTTCTCCAGCCTTTGTTCTCTAGGCTCAAGGCGCTCG 5412
1084 lUrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSerGly 1100
|||||
5413 AGTGGGTGAGCAGCGCGTGGAGCAGCTCCGGAGGCTACTACGGA 5462
1101 GlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTyrVa 1117
|||||
5463 GGAGACTTGCAAGTTCCACATCGCTTCCAAAGTTGGAGATTCAGTCTATGT 5512
1117 lArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProTyrL 1134
|||||
5513 TAGACGCCACCGTCGAGAAACCTCGAGACTCGGTGGAGGACCTTATC 5562
1134 euValLeuLeuThrThrProThrAlaValLysValGluGlyIleProThr 1150
|||||
5563 TCGTACTTTTGACCACACCAACGCGCTGTGAAAGTCGAAGGAATCCCCACC 5612
1151 TrpIleHisAlaSerHisValLysProAlaProProAspSerGlyTr 1167
|||||
5613 TGGATTCATGTCATCCAGCTTAAAGCGGCGCCACTCCCGACTCGGGGTG 5662
1167 pArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgLeuV 1184
|||||
5663 GAGACCCGAAAGACTGAGATCCCTTAAAGCTTCGCTCCATCGCCTGG 5712
1184 alProTyrSerAsnAsnSerProGlyGln 1194
|||||
5713 TTCCCTTACTCTAACAAATAACTCCCGAGGCCAG 5744

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:AAV09699

seq_documentation_block:

ID AAV09699 standard; DNA; 8196 BP.

XX AAV09699;

AC AAV09699;

XX 19-MAY-1998 (first entry)

XX Porcine retrovirus DNA encoding, GAG, POL and ENV.

XX Porcine retrovirus; PoEV; POL protein; ENV protein; GAG protein;

XX vaccine; diagnosis; xenotransplantation; prophylactic; therapeutic; ds.

XX Porcine retrovirus.

XX Key Location/Qualifiers

FT CDS 576..2126

FT /*tag= a

FT /product= GAG protein

FT /note= "vireon core polypeptide"

FT 2143..5733

FT /*tag= b

FT /product= POL protein

FT /note= "polymerase peptide sequence as given in

FT CDS 5606..7576

FT /*tag= c

FT /product= ENV protein

FT /note= "envelope protein"

XX W09740167-AL.

XX 30-OCT-1997.

XX 18-APR-1997; 97WO-GB01087.

XX 10-FEB-1997; 97GB-0002668.

XX 19-APR-1996; 96GB-0008164.

XX (IMUT-) IMUTRAN LTD.

XX (QONE-) Q-ONE BIOTECH LTD.

XX Galbraith DN, Haworth C, Lees GM, Smith KT;

XX WPI; 1997-535851/49.

XX Polynucleotide encoding porcine retrovirus expression product -

XX useful to develop products for use in vaccines, diagnosis and

XX xeno-transplantation

XX Claim 4; Fig 2; 69pp; English.

XX This DNA sequence encodes the porcine retrovirus (PoEV) vireon core

XX polypeptide (GAG), polymerase (POL) and envelope (ENV) proteins. These

XX proteins can be used to develop viral vaccines, antisense nucleic acids,

XX ribozymes and other antiviral agents. They can also be used in

XX xeno-transplantation technology and as diagnostic tools.

XX Sequence 8196 BP; 2165 A; 2061 C; 2147 G; 1820 T; 3 other;

alignment_scores:

Quality: 6284.00 Length: 1194

Ratio: 5.272 Gaps: 0

Percent Similarity: 99.832 Percent Identity: 99.749

alignment_block:

US-09-171-553B-5 x AAV09699 ..

Align seg 1/1 to: AAV09699 from: 1 to: 8196

1 GlyArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17
1249 GGGAGACGGGTTCGGACCCCTCCCGAGGCCAGGGTTAAGGT 2198
17 lGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34
2199 GGAGGGCAACCACTTGAGTTCCCTGGTTGATACCGGAGCGGAGCATTCAG 2248
34 aLeuLeuGlnProLeuGlyLysLeuLysGluLysLysSerTrpValMet 50
2249 TGCTGTACACCAATAGGAACCTAAAGAAAAAATCTCGGGTGATG 2298
51 GlyAlaThrGlyGlnArgGlnTrpTrpThrThrArgArgThrValAs 67
2299 GGTGCCACAGGCAACGGCATATCCATGGACTACCCGAAGAACCGTTGA 2348
67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysp 84
2349 CTTGGGAGTGGGACGGGTAAACCACTCGTTCTGTCATCCCTGAGTGCC 2398
84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100
2399 CAGTACCCCTCTAGGTAGACATTACTGACCAAGATGGGAGCTCAAAAT 2448
101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117
2449 TCTTTTGAACCAAGCAACCAAGAGTGTCTGTGAATAACAAACCCATCAC 2498
117 rValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnV 134
2499 TGTGTTGACCCCTCAATAGATGATGAATATCGACTATATCTCCCAAG 2548
134 aLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150
2549 TAAAGCCTGATCAAGATATACAGTCTCGTGGTGGAGCAGTTTCCCAAGCC 2598
151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProGlnVa 167
2599 TGGCAGAAACCGCAGGAGTGGTGGTGGCAAGCAAGTTCCCCACAGGT 2648
167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184
2649 TATTCACCTGAAGGCCAGTCTACACAGTATCAGTCAGACAGTACCCCT 2698
184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200
2699 TGAGTACAGAGGCTCGAGAGGAATTTGGCCGATGTTCAGAAATTAATC 2748
201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217
2749 CAACAGGGCATCTAGTTCCTGTCNAATCCCTTGGAAATCTCCCTGCT 2798
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234
2799 ACCGGTTAGGAAGCCTGGGACCAATGATTATCGACCAAGTACAGGACTTGA 2848
234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250
2849 GAGAGGTCAATAAAGGSGTCAGACATACACCAAGGTCGCCGACCCCT 2898
251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpThrValle 267
2899 TATAACCTCTTGAGCGCCCTCCCGCCTGAACGGAACTGGTACACAGTATT 2948
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284
2949 GGACTTAAAGATGCCCTTCTTCCTGAGATTACACCCCACTAGCCAAC 2998
284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
2999 CACTTTTGCTTCGATGAGAGATCCAGGTACGGGAAGAACCGGCGAG 3048
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317

3049 CTCACCTGGACCGACTGCCCAAGGGTTCAGAACTCCCGACCATCTT 3098
317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
3099 TGACGAAGCCCTACACAGGAGACTGGCCAACTTCAGGATCCCAACACCTC 3148
334 lNValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350
3149 AGGTGACCCCTCCCTCCAGTACGTGGATGACTGCTTCTGGCGGAGCCAC 3198
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuGluLeuSerAs 367
3199 AAACAGSACTGCTTAGAAGGTAGAGGCACTACTGCTGGAATTTGCTGA 3248
367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
3249 CTTAGGCTACAGAGCCTCTGCTAAGAGGCCAGATTTGCAGGAGAGG 3298
384 aThrTrpLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400
3299 TAACATACTTGGGGTACAGTTTCGGGGCGGCGAGCATGGCTGACGGAG 3348
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysGl 417
3349 GCACGGAAGAAACTCTAGTCCAGATACCGGCCCAACACAGCCAAACA 3398
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
3399 ACTGAGAGAGTTTTGGGGACAGCTGGATTTTGACACTGTGGATCCCGG 3448
434 lypPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450
3449 GGTTCGACCTTAGCAGCCCTCTACCCGCTAACCAAGAAAGAGG 3498
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467
3499 GGATTCCTCTGGGCTCTCAGCAGCAGAGAGCATTTGCTATCAAAA 3548
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484
3549 GGCCCTTGCTGAGCGCACCTGCTGCGCCCTCCCTGACGTAACTAAACCT 3598
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500
3599 TTACCCCTTTATGTGGATGAGCTAAGGAGTAGCCCGAGGAGTTTAAAC 3648
501 GlnThrLeuGlyProTrpArgProValAlaTyrLeuSerLysLysLe 517
3649 CAAACCTTAGGACCTAGGAGAGACCTGTGCTACCTGTCAAGAAGCT 3698
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaIav 534
3699 TGATCCTGTAGCCAGTGGTTGGCCGCTATGCTGAAGGCTATCGCAGCTG 3748
534 aIaIaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550
3749 TGGCCATACCTGTCAAGGAGCTGACAAATGACTTTGGGACAGATATA 3798
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567
3799 ACTGTAATAGCCCCCATGCATTGGAGAACATCGTTCCGCGAGCCCCAGA 3848
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeu 584
3849 CCGATGGATGACCAACGCCCGCATGACCCACTATCAAAAGCTTGCTTCTCA 3898
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600
3899 CAGAGAGGTCACCTTTCCGCTCCAGCGCTCTCAACCTGCCACTCTT 3948
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617

3949 CTGCGCTGAAGAGACTGTGTAACCACTGACTCATGATTGCCATCAACTATT 3998
 617 uileGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634
 3999 GATTGAGGAGACTGGGGTCCGCAAGGACCTTACAGACATACCGCTGACTG 4048
 634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650
 4049 GAGAAAGTCTAACCTGGTTCTACTGACGCAAGCAGCTATGTGTGGTGAAGGT 4098
 651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667
 4099 AAGAGGATGGCTGGGGCGGCGAGTGGTGACGGGACCCGACAGATCTGGGC 4148
 667 sSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684
 4149 CAGCAGCCTGCCGGAAGGAACCTTACGCCAANAGGCTGAGCTCATGGCCC 4198
 684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700
 4199 TCACGCAAGCTTTCGGGCTGGCGGCAAGGGAATCCATAAACATTTATACG 4248
 701 AspSerArgTyrAlaPheAlaHisValHisGlyAlaIleTyrLy 717
 4249 GACAGCAGTATGCTTTGGGACTGCACACGTACACGGGGCCATCTATAA 4298
 717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734
 4299 ACAAGGGGGTGTCTTACCTCAGCAGGGGAGGAAATAAAGAACAAAGAGG 4348
 734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750
 4349 AAATCTTAAGCCTATTAGAAGCCTTACATTTGCCAAAAGGCTAGCTATT 4398
 751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767
 4399 ATACACTGTCTCTGGACATCAGAAAGCCAAAGATCTCATATCTAGAGGGA 4448
 767 nGlnMetAlaAspArgValAlaLysGlnAlaGlnAlaValAsnLeuL 784
 4449 CCAGATGGCTGACCGGGTTGCCAGCAGCGACGCCAGCTGTTAACCTTC 4498
 784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800
 4499 TGCCTATAATAAGAACGCCCAAGCCCAAGAACCCAGACAGCTACAC 4548
 801 LeuGluAspTrpGlnIleLysLysIleAspGlnPheSerGluThrPr 817
 4549 CTAGAAAGACTGCCAAGAGATAAAGAGATAGACCACTTCTCTGAGACTCC 4598
 817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834
 4599 GGAGGGGACCTGCTATACCTCATATGGGAAGGAAATCCTGCCCAAAAG 4648
 834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850
 4649 AAGGGTTAGATATATCTCAACAGATACATCGTCTAACCCCACTAGGAAC 4698
 851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867
 4699 AAACACCTGCACGACTGGTCAGAACATCCCTTATCATGTTCTCAGGCT 4748
 867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884
 4749 ACCAGGAGTGGCTGACTCGGTGGTCAAAACATTTGTGTGCCCTGCCAGCTG 4798
 884 aAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900
 4799 TTAATGCTTAATCCTTCCAGAATACCTCCAGGAAGAGACTAAGGGGAAGC 4848
 901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917
 4849 CACCAGGCGCTCACTGGGAAGTGGACTTCACCTAGAGTAAAGCGGCTAA 4898

917 sTyrGlyAsnLysTyrIleLeuValPheValAspThrPheSerGlyTrpV 934
 4899 ATACGGAACAATAATATCTATTGTTTTTGTAGACACCTTTTCAGGATGG 4948
 934 alGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950
 4949 TAGAGGCTTATCTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAGAAA 4998
 951 IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967
 4999 ATACTGGAGGAAATTTTTCCAGATTGGATACCTAAGGTAATAGGGTC 5048
 967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984
 5049 AGACAATGTTCCAGCTTTCGTTGCCAGGTAAAGTCAGGACTGGCCAAGA 5098
 984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000
 5099 TATTGGGGATTGATTGGAACATCATTTGTATACAGACCCCAAGCTCA 5148
 1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017
 5149 GGCAGGTAGAGAGGATGAATAGAACCATTAAGAGACCTTACCAAAAT 5198
 1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034
 5199 GACCACAGACTGGCATTAATGATTGGATGGCTCTCTGCCCTTTGTGTC 5248
 1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050
 5249 TTTTGGGTGAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATGAA 5298
 1051 LeuLeuTyrGlyGlyProProLeuAlaGluIleAlaPheAlaHisSe 1067
 5299 TTGCTCTACGGGGGACCCCGCGTGGCAGAAATTCCTTTGCACATAG 5348
 1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084
 5349 TGCTGATGTCTGCTTTCCAGCCTTTGTCTTAGGCTCAAGGCGCTCG 5398
 1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSerGly 1100
 5399 AGTGGGTGAGGACGAGCGTGGAGCAGCTCCGGAGGCGCTACTCAGGA 5448
 1101 GlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTyrVa 1117
 5449 GGAGACTTGCAAGTTCCACATCGCTTCCAAGTTGGAGATTCAGTCTATGT 5498
 1117 lArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProTyrL 1134
 5499 TAGACGCCACCGTGCAGGAAACCTCGAGACTCGGTGGAGGGACCTTATC 5548
 1134 euValLeuLeuThrThrProThrAlaValLysValGluGlyIleProThr 1150
 5549 TCGTACTTTTGACCACACACCGCTGTGAAGTCGAGGAATCCCCACC 5598
 1151 TrpIleHisAlaSerHisValLysProAlaProProAspSerGlyTr 1167
 5599 TGGATCCATGCATCCCAACGTTAAGCYGGCGCCACCTCCCGACTCGGGGTG 5648
 1167 pArgAlaGluLysThrCluAsnProLeuLysLeuArgLeuHisArgLeuV 1184
 5649 GAGAGCGGAAAGACTGAGAAATCCCTTAAAGCTTCGCTCCATCGCCTGG 5698
 1184 alProTyrSerAsnAsnSerProGlyGln 1194
 5699 TTCTTACTTAACATTAACCTCCCGAGGCCAG 5730

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT: AAT74884
 seq_documentation_block:
 ID AAT74884 standard; cdna; 7892 bp.

XX

AAT74884;

XX

09-FEB-1998 (first entry)

XX

Miniature swine retrovirus cDNA.

XX

Retrovirus; porcine; GAG protein; POL protein; ENV protein;
 xenotransplantation; infectious; provirus; organ transplant; donor;
 activated virus; PCR; ss.

XX

Porcine retrovirus.

XX

Key Location/Qualifiers

FH

CDS 585..2159

FT

/*tag= a

FT

mat_peptide 585..2156

FT

/*tag= b

FT

/*note= "putative GAG protein"

FT

CDS 2307..5744

FT

/*tag= c

FT

mat_peptide 2307..5741

FT

/*tag= d

FT

/*note= "putative POL protein"

FT

CDS 5620..7536

FT

/*tag= e

FT

mat_peptide 5620..7533

FT

/*tag= f

FT

/*note= "putative ENV protein"

FT

W09721836-A1.

XX

19-JUN-1997.

XX

13-DEC-1996; 96WO-US19680.

XX

14-DEC-1995; 95US-0572645.

XX

(GEO) GEN HOSPITAL CORP.

XX

Fishman JA;

XX

WPI: 1997-32804/30.

XX

P-PSDB; AAW32096-W32098.

XX

New nucleic acid from porcine retroviruses - used for detecting
 viruses in transplant or other tissue and for assessing risk of
 transmitting infection to graft recipient

XX

Claim 22; Fig 3; 128pp; English.

XX

This cDNA sequence represents a porcine retrovirus from miniature swine
 containing the putative coding regions for viral GAG, POL and ENV
 proteins. This sequence and PCR fragments generated from it
 (see AAT74812-74882) can be used to screen organs for the presence of
 porcine retroviruses prior to xenotransplantation. Transplantation can
 increase the likelihood of retroviral activation if intact and
 infectious proviruses are present. The porcine retroviral sequence can be
 used to generate probes to determine the level (e.g. copy number) of
 intact (i.e. potentially replicating) porcine provirus sequences in a
 strain of xenograft transplantation donors. It can be used to detect
 mutations, genetic lesions or viral recombinants and also to determine
 the histological localisation of activated retrovirus. Using Polymerase
 Chain Reaction DNA Quantitation (PQD) on blood mononuclear cells
 infectivity titration and susceptibility testing can be performed.
 Ultimately animal donors without intact porcine retroviral sequences or a
 lower copy number of viral elements could be selected.

XX

Sequence 7892 BP; 2191 A; 1915 C; 1980 G; 1806 T; 0 other;

XX

alignment_scores:

Quality: 6099.50

Length: 1191

XX

Percent Similarity: 98.992

Ratio: 5.173

Gaps: 1

Percent Identity: 96.893

alignment_block:

US-09-171-553B-5 x AAT74884

..

Align seg 1/1 to: AAT74884 from: 1 to: 7892

1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17

|||||

2160 GGGAGACGGGGTTCGGACCCCTCCCGAGCCAGGGTAACCTTTGAAGT 2209

|||||

17 lGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34

|||||

2210 GGAGGGGCAACCACTTGTCTGTTGATACCGAGCGAACAACATTGAG 2259

|||||

34 allLeuLeuGlnProLeuGlyLysLeuLysGluLysLysSerTrpValMet 50

|||||

2260 TGCTACTACAGCCATTAGGAAACATAAAGATAAAAAATCCTGGGTGATG 2309

|||||

51 GlyAlaThrGlyGlnArgGlnTyrProTrpThrArgArgThrValAs 67

|||||

2310 GGTGCCACAGGGCAACACAGTATCCATGGACTACCCGGAAGAACAGTTGA 2359

|||||

67 pLeuGlyValGlyArgValThrHisSerPheLeuValTleProGluCysP 84

|||||

2360 CTTGGGAGTGGGACGGGTAACCCACTCGTTCTGTGTCATACCTGAGTGCC 2409

|||||

84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100

|||||

2410 CAGCACCCCTCTTAGGTAGAGACTTATTGACCAAGATGGAGCACAAATT 2459

|||||

101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117

|||||

2460 TCTTTTGAACAAGGGAAACCAAGAGTGTCTGCAAAATAACAACACCTATCAC 2509

|||||

117 rValLeuThrLeuGlnLeuAspGluTyrArgLeuTyrSerProGlnV 134

|||||

2510 TGTGTGACCCCTCCCAATTAGATGAGCAATATCGACTATCTCTCCCTAG 2559

|||||

134 allLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150

|||||

2560 TAAAGCCTGATCAAAATATACAAATCTCGTTGGAACAGTTTCCCAAGCC 2609

|||||

151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProGlnVa 167

|||||

2610 TGGGCAGAAACCGCAGGGATGGGTTGGCAAGCAAGTTCCCCCAAGT 2659

|||||

167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184

|||||

2660 TATTCACCTGAGGCCAGTGCACACCACTGTCTAGTCAGACAGTACCCT 2709

|||||

184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeu 200

|||||

2710 TCAGTAAAGAAGCTCAAGAGGAATTCGGCGCATGTCCAAAGATTAATC 2759

|||||

201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeu 217

|||||

2760 CAACAGGCGATCTCTAGTCTCTCCCAATCTCCCTGGAATCTCTCCCTGCT 2809

|||||

217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeu 234

|||||

2810 ACCGGTTAGAAGCTGGGACTAATGACTATCGACACAGTACAGGACTTGA 2859

|||||

234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250

|||||

2860 GAGAGGTCAATAAAGCGGTGCAGGATATACACCAACAGTCCCGAACCT 2909

|||||

251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrVal 267

|||||

2910 TATAACCTCTGTGTCTCTCCCAACCGGAGCTGGTATACAGTATT 2959

|||||

267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284

|||||
2960 GGACTTAAGGATGCTTCTTCTGCCTCAGATTACACCCCACTAGCCAAAC 3009
284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
|||||
3010 CACTTTTTCCTTCCGAATGGAGAGATCCAGGTAGCGGAAGAACCGGGCAG 3059
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317
|||||
3060 CTCACCTGGACCGGACTGCCCAAGGGTTCAAGAACTCCCGACCACTT 3109
317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
|||||
3110 TGACCAAGCCCTACACAGAGACCTGGCCCAACTTCAGGATCCACACCCCTC 3159
334 InValThrLeuLeuGlnTrpValAspAspLeuLeuAlaGlyAlaThr 350
|||||
3160 AGGTGACCCCTCCTCAGTACGTGGATGACCTGCTTCTGGCGGGAGCCACC 3209
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAs 367
|||||
3210 AACAGGAGCTGCTTAAAGGCACGAAGGCACACTACTGCTGGAAATGCTGCA 3259
367 pLeuGlyTrpArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
|||||
3260 CCTAGGCTACAGAGCCTCTGCTTAAAGGCCCGCAGATTTGCAGAGAGAGG 3309
384 alThrTrpLeuGlyTrpSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400
|||||
3310 TAACATACCTGGGGTACAGTTTGGCGGACGGGCAGCGATGGCTGACGGAG 3359
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG 417
|||||
3360 GCACGGAAGAAACTGTAGTCCAGATACCGGCCCCCAACACAGCCAAACA 3409
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
|||||
3410 AATGAGAGAGTTTGGGGACAGCTGGATTTTGCAGACTGTGGATCCCGG 3459
434 lyPheAlaThrLeuAlaAlaProLeuTrpProLeuThrLysLysGly 450
|||||
3460 GGTTCGACCTTAGCAGCCCACTCTACCGCTAACCAAGAAAGGG 3509
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467
|||||
3510 GAATTCCTCGGCTCCTCAGCAGCACCAGAAAGGCAATTTGATGCTATCAAAA 3559
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProP 484
|||||
3560 GGCCCTTGCTGAGGGCACCTGCTTGGCCCTCCTGACGTAACCTAACCCCT 3609
484 heThrLeuTrpValAspGluArgLysGlyValAlaArgGlyValLeuThr 500
|||||
3610 TPACCCCTTATGATGGATGAGCGTAAGGGAGTAGCCCGGGAGTTTAAACC 3659
501 GlnThrLeuGlyProTrpArgProValAlaTyThrLeuSerLysLysLe 517
|||||
3660 CAAACCCCTAGGACCATGGAGAGACCTGTCGCTACCTGCTCAAGAAGCT 3709
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaAlaV 534
|||||
3710 CGATCTGTAGCCAGTGGTGGGCCCATATGCTTGAAGGCTATCGCAGCTG 3759
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550
|||||
3760 TGCCCATCTGGTCAAGGAGCGCTGACAAATTTGACTTTGGACAGAAATA 3809
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567
|||||
3810 ACTGTAATAGCCCCCATGTCATTGGAGAACATCGTTTCGCGAGCCCCCAGA 3859
567 pArgTrpMetThrAsnAlaArgMetThrHisLysTrpGlnSerLeuLeuLeu 584
|||||

```
884 alaAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900
|||||
4810 TTAATGCTAATCTTCCAGAACTCCAGGGAAGAGACTAAGGGCAAGC 4859
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaIy 917
|||||
4860 CACCCAGGCGCTCAGTGGGAAGTGGACCTTCACTGAGGTAAAGCGGCTAA 4909
917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934
|||||
4910 ATACGGNAACAATACTATTTGTTTGTAGACACCTTTTCAGGATGGG 4959
934 alGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950
|||||
4960 TAGAGGCTTATCTACTAAGAAGAGACTTCAACCGTGTGGCTAAAAAAA 5009
951 IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967
|||||
5010 ATACTGGAAGAAATTTTCCAGATTGGAATACCTTAAGTAAATAGGGTC 5059
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984
|||||
5060 AGACAATGTCCAGCTTTTGTGCCAGGTAACTCAGGACCTGGCCCAAGA 5109
984 leLeuGlyIleAspThrLysLeuHisCysAlaTyrArgProGlnSerSer 1000
|||||
5110 TATTGGGATTGATTGGAACCTGCAATTGTGCATACAGACCCCAAGCTCA 5159
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017
|||||
5160 GGACAGGTAGAGAGGATGAATAGAACCATTAAGAGACCCCTTACTAAATT 5209
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034
|||||
5210 GACCGCGAGACTGGCGTTAATGATTGATAGTCTCTCTGCCCTTTGGCC 5259
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050
|||||
5260 TTTTGGGTTAGGAACACCCCTGCACAGTTTGGCTGACCCCTATGAA 5309
1051 LeuLeuTyrGlyGlyProProProLeuAlaGluIleAlaPheAlaHisSe 1067
|||||
5310 TTACTCTACGGGGGACCCCCCATTTGGTAGAAATTTGCTGTACATAG 5359
1067 rAlaAspValLeuLeuSerGlnProPhePheSerArgLeuLysAlaLeuG 1084
|||||
5360 TGCTGACGTCGCTTTCCAGCCCTTGTCTCTAGGCTCAAGGCACCTTG 5409
1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSer... 1099
|||||
5410 ACTGGTGAGACACACGAGCGTGGAGGCAACTCCCGGAGGCTTACTCAG 5459
1100 GlyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTy 1116
|||||
5460 GGAGGAGACTTCAGATCCCATCGTTTCCAAAGTGGGAGATTCAGTCTA 5509
1116 rValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProT 1133
|||||
5510 CATTAGACCCACCTCGCAGGAAACCTCGAGACTCGGTGGAGGCGCCCTT 5559
1133 yLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIlePro 1149
|||||
5560 ATCTCGTACTTTTGACCAACACACGCGCTGTGAAAGTCGAAGTAATCTCC 5609
1150 ThrTrpIleHisAlaSerHisValLysProAlaProProAspSerGly 1166
|||||
5610 ACCTGGATCCATGATCCACAGTTAAACCGGCGCCACCTCCCGATTCGGG 5659
1166 yTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgL 1183
|||||
5660 GTGGAAACCCGAAAGACTGAATATCCCTTTAAGCTTCGCGCTCCATCGCG 5709
```

```
1183 euValProTyrSerAsnAsnAsn 1190
:::|||||
5710 TGGTTCTTACTCTGTCAATAAC 5732
```

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF77727

seq_documentation_block:

ID AAF77727 standard; DNA; 8132 BP.

XX AAF77727;

XX 23-MAY-2001 (first entry)

XX Nucleotide sequence of a retrovirus found in miniature swine.

XX Retrovirus; graft transplantation; xenotransplantation; miniature swine;

XX ss.

XX Unidentified.

XX US6190861-B1.

XX 20-FEB-2001.

XX 13-DEC-1996; 96US-0766528.

XX 14-DEC-1995; 95US-0572645.

XX (GEO) GEN HOSPITAL CORP.

XX Fishman JA;

XX WPI; 2001-256211/26.

XX P-PSDB; AAB73285, AAB73286, AAB73287.

XX Assessing risk of endogenous retroviruses in clinical practice and in

XX or miniature swine retroviral genome

XX Claim 1; Fig 3; 127pp; English.

XX The present invention relates to a method for screening a cell or tissue

XX for the presence or expression of a retrovirus (RV), comprising

XX contacting a target nucleic acid from the cell or tissue with a second

XX nucleic acid from the present invention (e.g. the present sequence or a

XX fragment thereof). The method is useful for RV detection and to assess

XX graft transplantation risk. Screening of animals allows the elimination

XX of donors with active replication of known viruses. Inactive proviruses

XX can be detected and inactivated, allowing identification and elimination

XX of potential human pathogens derived from swine in a manner not possible

XX in the outbred human organ donor population and is important to the

XX development of human xenotransplantation.

XX SQ Sequence 8132 BP; 2248 A; 1977 C; 2037 G; 1870 T; 0 other;

alignment_scores:

Quality: 6099.50 Length: 1191

Ratio: 5.173 Gaps: 1

Percent similarity: 98.992 Percent Identity: 96.893

alignment_block:

US-09-171-553B-5 x AAF77727 ..

Align seg 1/1 to: AAF77727 from: 1 to: 8132

1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17

1183 euValProTyrSerAsnAsnAsn 1190

2160 GGGAGACGGGGTTCGGACCCCTCCCGCAGCCAGGGTAACCTTGAAGGT 2209

17 lGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisServ 34

2210 GGAGGGCAACCAAGTTGAGTTCTGTTGATACCGGAGCGAAACATTCAG 2259

34 aLeuLeuGlnProLeuGlyLysLeuLysGluLysSerTrpValMet 50
|||||
2260 TGCTACTACAGCCATTAGGAAACTAAAAAGATAAAAAATCCTGGTGTATG 2309
51 GlyAlaThrGlyGlnArgGlnTyrProThrThrArgArgThrValAs 67
|||||
2310 GGTGCACAGGGCAACACACATATCCATGGACTACCCGAAACAGATTGA 2359
67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysp 84
|||||
2360 CTGGGAGTGGAGCGGTAACCCACTGCTTCTGTCATACCTGAGTGCC 2409
84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100
|||||
2410 CAGCACCCCTCTTAGGTAGACATTATTGACCAAGATGGGAGCACAAAT 2459
101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117
|||||
2460 TCTTTTGAACAGGGAACCAAGAGTGTCTCAAAATACAAACATATCAC 2509
117 rValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnV 134
|||||
2510 TGTGTTGACCCCTCCNAATTAGATGACGAATATCGACTATCTCTCCCTAG 2559
134 aLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150
|||||
2560 TAAAGCCTGATCAAAATATACAATCTGTTGGAACAGTTTCCCCAAAGCC 2609
151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnVa 167
|||||
2610 TGGGAGAAACCGCAGGAGTGGTTTGGCAAGCAAGTTTCCCCACAAAGT 2659
167 IleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184
|||||
2660 TATTCACTGAAGGCCAGTGCACACACAGTGTCTAGTCAGACAGTACCCCT 2709
184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200
|||||
2710 TGAGTAAAGAAAGCTCAAGAAAGAAATTCGGCGCATGTCTCAAGATTAATC 2759
201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217
|||||
2760 CAACAGGCATCTAGTTCCTCCCAATCTCCCTGGAAATCTCCCTGCT 2809
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234
|||||
2810 ACCGGTTAGAAAGCCTGGGACTAATGACTATCGACCATCGACAGACATTGA 2859
234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250
|||||
2860 GAGAGGTCAATAAAGGGGTGAGGATATACACCCCAACAGTCCCGAACCCCT 2909
251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValIle 267
|||||
2910 TATAACCTCTCTGTGCTCTCCACCCCAAGGAGTGGTATACAGATT 2959
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284
|||||
2960 GGACTTAAAGGATGCTTCTTCTGCTGAGATTACACCCCACTAGGCCAAC 3009
284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
|||||
3010 CACTTTTTCCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCGGGCAG 3059
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317
|||||
3060 CTCACCTGGACCCGACTGCCCAAGGGTTCAAGAACTCCCGACCACTT 3109
317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
|||||
3110 TGACGAAGCCCTTACACAGAGACCTGGCCAACTTCAGGATCCAAACACCCCTC 3159

334 InValThrLeuLeuGlnTyrValAspLeuLeuAlaGlyAlaThr 350
|||||
3160 AGGTGACCTCTCCAGTACGTGGATGACCTGCTTTCGGCGGAGCCACC 3209
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAs 367
|||||
3210 AATCAGGACTGCTTAGAAGGACAGGACACTACTGCTGGAATTTGCTGA 3259
367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgGluV 384
|||||
3260 CCTAGGCTACAGAGCCTGCTTAAGAAGGCCAGATTTCAGGAGAGAGG 3309
384 alThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400
|||||
3310 TAACTACTTGGGTACAGTTTTCGGGACGCGCAGCATGGCTGACGGAG 3359
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG 417
|||||
3360 GCACGGAAGAAACTGTAGTCAGATACCGGCCCAACCCACAGCCAAACA 3409
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
|||||
3410 AATGAGAGAGTTTTGGGACACAGCTGGATTTCGACACTGTGGATCCCGG 3459
434 lyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450
|||||
3460 GGTTCGGACCTTAGCAGGCCCTCTACCCGCTAACCAAGAAAGAGG 3509
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467
|||||
3510 GAATTCCTCTGGGCTCTTGAGCACCAGAGAGCATTTGATGCTATCAAAA 3559
467 saLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484
|||||
3560 GGCCCTGCTGAGCGCACCTGCTCTGGCCCTCCCTGACCTAACTAAACCT 3609
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500
|||||
3610 TTACCTTTATGTGGATGAGCGTAAGGAGTAGCCCGGGAGTTTAAACC 3659
501 GlnThrLeuGlyProTrpArgProValAlaTyrLeuSerLysLysLe 517
|||||
3660 CAACCCCTAGGACCATGGAGAACCTGTGCGCTACCTGTCAAAAGAGCT 3709
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaAlaV 534
|||||
3710 CGATCTCTAGCCAGTGGTTGGCCCATATGCTGAAGGCTATCGCAGCTG 3759
534 alaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550
|||||
3760 TGGCCATCTGGTCAAGGACCTGACAAATTTGACTTTGGGACAGAAATA 3809
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567
|||||
3810 ACTGTAATAGCCCCCATGCTTGGAGAACATCGTTTCGGCAGGCCCCAGA 3859
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuT 584
|||||
3860 CCGATGGATGACCAACGCGCATGACCCACTATCAAGGCTGCTCTCTCA 3909
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600
|||||
3910 CAGAGAGGTCAGTTCGCTCCACAGCCGCTCTCAACCTGCCACTCTT 3959
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617
|||||
3960 CTCCTGAAGAGACTGATGAACCCAGTCACTCATGATTGCCATCAACTAT 4009
617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634
|||||
4010 GATTGAGGAGACTGGGGTCGCAAGGACCTTACAGACATACCCGCTGACTG 4059
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650

||||| 4060 GAGAAAGTCTAACCTGGTTCCTACTGACGGAACAGCAGTATGTGGTGAAGGT 4109
651 LysArgMetAlaGlyAlaValAlaValAspGlyThrArgThrIleTrpAl 667
||||| 4110 AAGAGGATGGTGGGGCGCGGTGGTGGACGGGACCCGCACGATCTGGGC 4159
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684
||||| 4160 CAGCAGCCTGCCGGAAGAACTTCAGCACAAAAGGCTCAGCTCATGGCCC 4209
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700
||||| 4210 TCAGCAAGCTTTCGGCGTGGCCGAAGGGAATCCATAAATTTATACG 4259
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717
||||| 4260 CACAGCAGGTATGCTTTGGCAGTCGCACAGTACATGGGGCCATCTATAA 4309
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734
||||| 4310 ACAAGGGGTTGCTTACCTCAGCAGGAGGGGAATAAAGAACAAAGAGG 4359
734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750
||||| 4360 AAATCTAAGGCTATTAGAAGCCGTACATTTACCAAAAAGGCTAGCTATT 4409
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767
||||| 4410 ATACACTGCTCGACATCAGAAAGCTAAGATCTCATATCCAGAGGAAA 4459
767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784
||||| 4460 CCAGATGGCTGACCGGGTTGCCAAGCAGGAGCCAGCCAGGGTTAAACCTTC 4509
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800
||||| 4510 TGCCTATAATAGAAATGCCCAAGCCCGCAGAACCCAGACGACAGTACACC 4559
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817
||||| 4560 CTAGACACTGGCAAGAGATAAAGAGATAGACCAGTTCTCTGAGACTCC 4609
817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834
||||| 4610 GGAAGGGACCTGCTATACCTCAGATGGGAAGAAATCCTGCCCCACAAAG 4659
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850
||||| 4660 AAGGTTTAGAATATGTCACAGATACATCGTCTAACCACCTAGGAACCT 4709
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867
||||| 4710 AAACACCTGCAGCAGTTGGTCAGAACATCCCTTATCATGTTCTGAGGCT 4759
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884
||||| 4760 ACCAGGAGTGGCTGACTCGGTGGTCAAAACATTTGTGCCCTGCCAGCTGG 4809
884 alaAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900
||||| 4810 TTAATGCTAATCTTCAGAAATGCCCTCCAGGGAAGAGACTAAGGGGAAGC 4859
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917
||||| 4860 CACCCAGGCCCTCACTGGGAAGTGGACTTCACCTGAGGTAAAGCCGGCTAA 4909
917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934
||||| 4910 ATACGGAACAAATACTATTGTTTTTGTAGACACCTTTTCAGGATGGG 4959
934 alGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950
|||||

4960 TAGAGGCTTATCTCTACTACAAAAGAGACTTCAACCCCTGGTGGCTAAAAA 5009
951 IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967
||||| 5010 ATACTGGAAGAAATTTTCCAGATTGGAAATACCTAAGGTAATAGGGTC 5059
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984
||||| 5060 AGACAAATGGTCCAGCTTTTGTGCCCCAGGTAAAGTCAGGAGCTGGCCAAG 5109
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000
||||| 5110 TATTGGGGATTGATTGGAACCTGATTGTGCATACAGACCCCAAGCTCA 5159
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017
||||| 5160 GGACAGGTAGAGAGATGAATAGAACCATTAAGAGACCCCTTACTAAAT 5209
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034
||||| 5210 GACCGCGGAGACTGGCGTTAATGATTGGATAGCTCTCTCTGCTTGTGTC 5259
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050
||||| 5260 TTTTATAGGTTAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATGAA 5309
1051 LeuLeuTyrGlyGlyProProLeuAlaGluIleAlaPheAlaHisSe 1067
||||| 5310 TTACTCTACGGGGACCCCCCATTTGGTAGAAATTTGCTTGTACATAG 5359
1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084
||||| 5360 TGCTGACGTGCTGCTTCCAGCCTTTGTCTCTAGGCTCAAGGCACCTG 5409
1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSer... 1099
||||| 5410 AGTGGGTGAGACAACGAGCGTGGAGCAACTCCGGGAGGCTTACTCAGGA 5459
1100 GlyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTy 1116
||||| 5460 GGAGGAGACTTCAGATCCCATCCATCGTTTCCAAGTGGGAGATTCACTTA 5509
1116 rValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpTyrGlyProT 1133
||||| 5510 CGTTAGACCCACCGGTGCGAGGAAACCTCGAGACTCGGTGGAAGGCCCTT 5559
1133 yrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIlePro 1149
||||| 5560 ATCTCGTACTTTTGACCACACACACGCTGTGAAAGTCGAAGAAATCTCC 5609
1150 ThrTrpIleHisAlaSerHisValLysProAlaProProAspSerGly 1166
||||| 5610 ACCTGGATCCATGCATCCACGTTAAACCGGGCCACCTCCCGATTCGGG 5659
1166 yTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgL 1183
||||| 5660 GTGGAAGCCGAAAGACCTGAAATCCCTTAAGCTTCGCTCCCTCATCGG 5709
1183 euValProTyrSerAsnAsnAsn 1190
:::||||| 5710 TGGTTCTTACTCTGTCAATAAC 5732

seq_name: /SIDS2/gcgdata/geneseq/NA2001.DAT: AAC67020

seq_documentation_block:

ID AAC67020 standard; DNA; 4402 BP.

XX AAC67020;

XX AC

XX DT 27-MAR-2001 (first entry)

XX PERV env protein coding sequence SEQ ID NO: 20.

XX

XX

KW Xenotransplantation; infectious agent; vaccine; ds.

OS Porcine endogenous retrovirus.

XX W0200071726-A1.

XX 30-NOV-2000.

XX 24-MAY-2000; 2000WO-US14296.

XX 24-MAY-1999; 99US-0135631.

XX (MAYO-) MAYO MEDICAL VENTURES.

XX Federspiel MJ;

XX WPT; 2001-032041/04.

XX Inhibiting or preventing infectious agent transmission in mammalian
PT transplant recipients, by introducing recombinant DNA comprising DNA
PT encoding extracellular proteins of the agent into donor cells, such as
PT swine cells -

XX Claim 16; Page 105-106; 144pp; English.

XX The present invention provides a method to prevent the transmission of
CC infectious agents during xenotransplantation. This involves introducing
CC to donor swine cells a recombinant DNA encoding a peptide fragment from
CC the infectious agent, and then introducing these cells into the
CC transplant recipient.

XX SQ Sequence 4402 BP; 1259 A; 1085 C; 1111 G; 947 T; 0 other;

alignment_scores:

Quality: 6095.50 Length: 1192
Ratio: 5.174 Gaps: 1
Percent Similarity: 98.826 Percent Identity: 96.980

alignment_block:

US-09-171-553b-5 x AAC67020 ..

Align seg 1/1 to: AAC67020 from: 1 to: 4402

1 GlyArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17
224 GGAAGACGGGGTTCGGACCCCTCCCGGAGCCAGGGTAACCTTGAAGGT 273
17 lGluGlnProValGluPheLeuValAspThrGlyAlaGluHisServ 34
274 GGAGGGGCAACAGTTGAGTTCCTGGTTGATACCGGAGCGAAACATTTCAG 323
34 alLeuLeuGlnProLeuGlyLysLeuLysGluLysLysSerTrpValMet 50
324 TGCTACTACAGCCATTAGGAAACTTAAAGATATAAATAATCCCTGGGTGATG 373
51 GlyAlaThrGlyGlnArgGlnTyrProThrThrArgArgThrValAs 67
374 GTGCCACAGGCGCAACACAGTATCCATGGACTCCCGAAGAACAGTTGA 423
67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysP 84
424 CTTGGAGTGGGAGGGTAAACCACTCGCTTCTGGTTCATACCTGAGTGCC 473
84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100
474 CAGACCCCTCTAGTAGAGACTTATTGACCACAGATGGGAGCACAAAT 523
101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117
524 TCTTTTGAACAGGGAACCAAGAGTCTGCAAAATAACAAACCTATCAC 573
117 rValLeuThrLeuGlnLeuAspGluTyrArgLeuLysSerProGlnV 134

574 TGTGTTGACCTCCAATTAGATGACGAATATCGACTATATCTCTCCCTAG 623
134 allYsProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150
624 TAAAGCCTGATCAAAATATACAAATCTGGTTGGAACAGTTTCCCCAAGCC 673
151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnVa 167
674 TGGGCAGAAACCGCAGGATGGTTTGGCAAGCAAGTTCCCCACAAGT 723
167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184
724 TATTCAACTGAAGGCCAGTGCACACAGTGTCTGTCAGTCAGACAGTACCCCT 773
184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200
774 TGAGTAAGAGCTCAAGAGGAATTCGGCCGCATGTCACCAAGATTAATC 823
201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217
824 CAACAGGGCATCTAGTTCCTGTCCAATCTCCCTGGAATACTCCCTGCT 873
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234
874 ACCGGTTAGAAAGCCTGGGACTAATGACTATCGACCAGTACAGGACTTGA 923
234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250
924 GAGAGTCAATAAACGGTGCAGGATATACACCAACAGTCCCGAACCCCT 973
251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValle 267
974 TATAACCTCTGTCTCTCCACCCCAACGAGGCTGTATACAGTATT 1023
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284
1024 GGACTTAAAGGATGCCCTTTCTGCTGAGATTACACCCCACTAGCCAAC 1073
284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
1074 CACTTTTTCCTTCGATGGAGATCCAGTACGGGAGAACCGGCGAG 1123
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317
1124 CTCACCTGGACCCGACTGCCCCAAGGGTTCAAGAACTCCCGCCCATCTT 1173
317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
1174 TGCGAAGCCCTACACAGAGACCTGGCCAACTTCAGGATCCACACCCCTC 1223
334 lnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350
1224 AGGTGACCCCTCCAGTACGTGATGACCTGCTTCTGGCGGGAGCCACC 1273
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAs 367
1274 AAACAGACTGCTTAGAAGGCACGAGGCTACTGCTGGAATGTCTGA 1323
367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
1324 CCTAGGCTACAGACCTCTGCTAAGAGGCCAGATTTGCGAGGAGAGG 1373
384 alThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400
1374 TAACATACTTGGGGTACAGTTTGGCGACGCGGACGATGGCTGACGGAG 1423
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysGl 417
1424 GCAGGGAAGAAACTGTAGTCCAGATACCGGCCCAACACAGCCAAACA 1473
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434

1474 AGTGAGAGACTTTTTGGGGACAGCTGGATTTTGCAGACCTGGATCCCGG 1523
 434 lyPheAlaThrLeuAlaAlaProLeuThrProLeuThrLysGluLysGly 450
 1524 GTTTGGGACCTTAGCAGCCCCACTCTACCCACTAACCAAGAAAGGG 1573
 451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467
 1574 GAATCTCTCTGGGCTCTGAGCAGCAGAGGCAATTTGATGCTATCAAAA 1623
 467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysPro 484
 1624 GGCCCTCTGAGCGCACCTGCTTGCCCTCCCTGAGTAACTAAACCCCT 1673
 484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500
 1674 TTACCCCTTATGTGGATGAGCGTAAGGAGTAGCCCGGGAGTTTAAACC 1723
 501 GlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLe 517
 1724 CAAACTTAGGACCATGGAGGAGACCTGTTGCCTACCTGTCAAGAAGCT 1773
 517 uAspProValAlaSerClyTrpProValCysLeuLysAlaIleAlaLav 534
 1774 CGATCTGTAGCCAGTGGTTGGCCCTGATGCCCTGAAGGCTATCGCAGCTG 1823
 534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550
 1824 TGGCCATCTGGTCAAGGACCGTGACAAATTGACTTTGGGACAGAATAA 1873
 551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567
 1874 ACTGTAATAGCCCCCATGCTGGTGGAGAACATCGTTTCGGCAGCCCCAGA 1923
 567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuT 584
 1924 CCGATGATGACCAACGCCCGCATGACCCCACTATCAAGCCCTGCTCTCA 1973
 584 hrGluArgValThrPheAlaProAlaAlaLeuAsnProAlaThrLeu 600
 1974 CAGAGAGGTTCAGTTCGCTCCACAGCCGCTCTCAACCCCTGCCACTCTT 2023
 601 LeuProGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617
 2024 CTGCCTGAAGAGACTGATGAACCACTGACTCATGATTGCCATCAACTATT 2073
 617 uIleGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634
 2074 GATTGAGGAGACTGGGGTCGCGAAGGACCTTACAGACATACCCGCTGACTG 2123
 634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650
 2124 GAGAAGTGTAACTGCTGCTGCTGAGCGAAGCAGCTATGTGTGGGAAGGT 2173
 651 LysArgMetAlaGlyAlaValAlaValAspGlyThrArgThrIleTrpAl 667
 2174 AAGAGGATGCTGGGCGCGGGTGGTGGACGGGACCCGACGATCTGGGC 2223
 667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684
 2224 CAGCAGCCTGCGCGAAGGAACTTTCAGCACAAAAGGCTGAGCTCATGGCCC 2273
 684 euThrGlnAlaLeuArgLeuAlaGluLysSerIleAsnIleTyrThr 700
 2274 TCACGAAGCTTTGGCGCTGGCGAAGGGAATCCATAAACAATTTATACA 2323
 701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717
 2324 GACAGCAGGTATGCTTTGGGACTGCACAGCTACACGGGGCCATCTATAA 2373
 717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluC 734
 2374 GCAAAGGGGCTTCTTACCTCAGCAGGGAGGGAATAAAGAACAAAGAGG 2423

734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750
 2424 AAATTTCTAAGCCTATTAGAAAGCCTTACATTTGCCAAAAGGCTGATATT 2473
 751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767
 2474 ATACACTGTCTGGACATCAGAAAGCCAAAGATCCCATATCCAGAGGAA 2523
 767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784
 2524 CCAGATGGCTGACCGGGTTGCCAAGCAGCAGCCAGGGGTGTAACCTTC 2573
 784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800
 2574 TGCCCTATGATAGAAACACCCCAAGCCCAAGACCCCGGAGCAGTACACC 2623
 801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817
 2624 CTAGAAGACTGGCAAGAGATAAAAAGATAGACCAGTTCTCTGACACTCC 2673
 817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834
 2674 GGAAGGACCTGCTATACCTCAGATGGGAAGGAAATCCTGCCCCACAAG 2723
 834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850
 2724 AAGGGTTAGAAATATGTCCACAGATACATCGTCTAACCCACCTAGGAAC 2773
 851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867
 2774 AAACACCTGCACAGTGTGTCAGAACATCTCCTTATCATGTCTGAGGCT 2823
 867 uProGlyValAlaAspSerValLysHisCysValProCysGlnLeuV 884
 2824 ACCAGAGTGGCTGATTCGGTGGTCAAACTGTGTGCCCTGCCAGCTGG 2873
 884 alAsnAlaAsnProSerArgIleProGlyLysArgLeuArgGlySer 900
 2874 TTAATGCTAATCCTTCCAGATATCCTCCAGGAAGAGACTAAGGGGAAGC 2923
 901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917
 2924 CACCAGCGCGCTCACTGGCAAGTGGACTTCACTGAGGTAAGCCGGCTAA 2973
 917 styGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934
 2974 ATACGGAAACAAATATCTATTGGTTTTTTGTAGACACCTTTTCAGGATGG 3023
 934 alGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950
 3024 TAGAGCTTATCTCTACTAAGAAGAGACTTCAACCGTGGTGGCTAAGAAA 3073
 951 IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967
 3074 ATACTGGAGGAAATTTTCCAAGATTTGGAATACCTAAGGTAATAGGTC 3123
 967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984
 3124 AGACAATGCTCCAGCTTTCGTTGCCAGGTAAGTCAGGGAAGTGGCCAAAG 3173
 984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000
 3174 TATTTGGGATTTGATTGGAACTGTCATTTGCATACAGACCCCAAGCTCA 3223
 1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017
 3224 GGAAGGTAGAGAGGATGAATAGAACCATTAAAGAGACCCCTTACTAAAT 3273
 1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034
 3274 GACCGCGGAGACTGGCGTTAATGATTGGATAGTCTCTCCTGCGCTTGTGC 3323

1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050
 3324 TTTTATAGGTTTATAGAACACCCCTGGACAGTTTGGGCTGACCCCTATGAA 3373
 1051 LeuLeuTyrGlyGlyProProLeuAlaGluLeuAlaPheAlaHisSe 1067
 3374 TTACTCTACGGGGACCCCTCCCATTTGTAGAAATGCTTCCGTCACATAG 3423
 1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084
 3424 TGCTGACGTGCTGCTTCCAGCCCTTGTCTTCTAGGCTCAAGGCACCTG 3473
 1084 LuTrpValArgGlnAlaTrpLysGlnLeuArgGluAlaTyrSer... 1099
 3474 AGTGGGTGAGACAACGAGCGTGGAGGCAACTCCGGGAGGCTACTCAGGA 3523
 1100 GlyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTy 1116
 3524 GGAGGAGACTTGCAGATCCCATCGTTTCCAAAGTGGGAGATTCAGTCTA 3573
 1116 rValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProT 1133
 3574 CGTTAGAGCCACCGCTGAGAAACCTCGAGACTCGGTGGAGGGCCCTT 3623
 1133 yrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIlePro 1149
 3624 ATCAGTACTTTTACCACACCAACGCGTGTGAAAGTCGAAGGAATCTCC 3673
 1150 ThrTrpIleHisAlaSerHisValLysProAlaProProAspSerG1 1166
 3674 ACCTGGATTCATGATCCACGTTTAAGCGCGCCACCTCCCGATTCCGG 3723
 1166 yTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgL 1183
 3724 GTGGAACCCGAAAGACTGAAATCCCTTAAGCTTCGCTCCCATCGCG 3773
 1183 euValProTyrSerAsnAsnSer 1191
 3774 TGGTTCCTTACTCTGTCAATAACTCC 3799

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAC67019

seq_documentation_block:

ID AAC67019 standard; DNA; 7362 BP.

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX AC AAC67019;

XX

PS Claim 16; Page 101-104; 144pp; English.

XX

CC The present invention provides a method to prevent the transmission of
 CC infectious agents during xenotransplantation. This involves introducing
 CC to donor swine cells a recombinant DNA encoding a peptide fragment from
 CC the infectious agent, and then introducing these cells into the
 CC transplant recipient.

XX

SQ Sequence 7362 BP; 1997 A; 1821 C; 1881 G; 1663 T; 0 other;

alignment_scores:

Quality: 6095.50 Length: 1192

Ratio: 5.174 Gaps: 1

Percent Similarity: 98.826 Percent Identity: 96.980

alignment_block:

US-09-171-553B-5 x AAC67019 ..

Align seg 1/1 to: AAC67019 from: 1 to: 7362

1 GlyArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17
 2823 GGAAGACGGGTTCGGACCCCTCCCGAGCCCGAGGTAACTTTGAAGT 2872
 17 lGluGlyGlnProValIgluPheLeuValAspThrGlyAlaGluHisSerV 34
 2873 GGAGGGCAACCACTTCTGCTGATACCGGAGCGAAACATTCAG 2922
 34 alLeuLeuGlnProLeuGlyLysLeuLysGlyLysSerTrpValMet 50
 2923 TGTACTACAGCCATTAGGAAACTAAAGATAAAATCCCTGGGTGATG 2972
 51 GlyAlaThrGlyGlnArgGlnTyrProTrpThrArgArgThrValAs 67
 2973 GGTGCCACAGGGCAACACAGATATCCATGACTACCCGAGAACAGTTGA 3022
 67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysP 84
 3023 CTTGGAGTGGGACGGGTAAACCCACTCGTTCTGTGTCATACCTGAGTGC 3072
 84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100
 3073 CAGCACCCCTCTTAGGTAGAGACTTATTGACCAAGATGGGACACAAAT 3122
 101 SerPheGluGlnGlyArgProGluValSerValAsnLysProIleTh 117
 3123 TCTTTTGACACAGGGAACACAGAGTGTCTGCAATAACAAACCTATCAC 3172
 117 rValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnV 134
 3173 TGTGTTGACCCCTCCAATTAGATGACGAATATCGACTATATCTCTCCCTAG 3222
 134 alLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150
 3223 TAAAGCCTGATCAAAATATACAAATCTGTTGGACAGTTTCCCAAGCC 3272
 151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnVa 167
 3273 TGGGCAAGAACCGCAGGATGGTTTGGCAAGCAAGTTCCCCCAAGT 3322
 167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184
 3323 TATTCAACTGAAGGCGAGTCCACACAGTGTTCAGTCAGACAGTACCCCT 3372
 184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200
 3373 TGAGTAAAGAAAGCTCAAGAGGAATTCGGCCGATGTCCAAAGATTATC 3422
 201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217
 3423 CAACAGGGCATCTAGTCTCTGTCCAATCTCCCTGGGAATCTCCCTGCT 3472

217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234
|||||
3473 ACCGGTTAGAAAGCCTGGGACTATGACTATCCACCAGTACAGAGCTTGA 3522
234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250
|||||
3523 GAGAGGTCAATAAACGGGTGAGGATATACACCAACAGTCCCGAACCCCT 3572
251 TyrAsnLeuLeuSerAlaLeuProGluArgAsnTrpThrValle 267
|||||
3573 TATAACCTCTGTGTGCTCTCCACCCCAACGGAGGTGGTATACAGTAT 3622
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284
|||||
3623 GGACTTAAGAGGTGCTTTTCTGCTGAGATTACACCCCACTAGCCAAC 3672
284 roLeupheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
|||||
3673 CACTTTTGGCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCGGGCAG 3722
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317
|||||
3723 CTCACCTGGACCCGACTGCCCAAGGGTTCAAGAACTCCCGACCATCTT 3772
317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
|||||
3773 TGACGAAGCCCTACACAGAGACCTGCCCAACTTCAGAGTCCCAACACCCCTC 3822
334 lnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350
|||||
3823 AGGTGACCCCTCCFCCAGTACGTGGATGACCTGCTTCTGGCGGGAGCCACC 3872
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuLeuSerAs 367
|||||
3873 AAACAGACTGCTTAGAAGGCACCAAGGCACACTACTGCTGGAATGTCTGA 3922
367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
|||||
3923 CCTAGGCTACAGAGCCTCTGCTAAGAAGCCCAAGATTTGCAGAGAGAGG 3972
384 alThrTrpLysLeuGlyTyrSerLeuArgGlyClyGlnArgTrpLeuThrGlu 400
|||||
3973 TAACATACTTGGGTACAGTTTGGGGACGGCAGCGATGGCTGACGGAG 4022
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG 417
|||||
4023 GCACGAAGAAGAACTGTAGTCCAGATACCGGCCCAACACACAGCCAAACA 4072
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
|||||
4073 AGTGAGAGAGTTTGGGACACAGCTGGATTTGCAGACTGTGGATCCCGG 4122
434 lypPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGlyLysGly 450
|||||
4123 GGTTCGACCTTAGCAGCCCACTCTACCCCACTAACCAAGAAAGGAGG 4172
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLys 467
|||||
4173 GAATCTCTCTGGCTCCTGAGCACACAGAGGCATTTGATGCTATCAAAA 4222
467 salAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484
|||||
4223 GGCCCTGTGTAGCGCACCTGCTCTGGCCCTCCCTGACCTAACTAACCCCT 4272
484 heThrLeuTyrValAspGluArgLysGlyValAlaAlaArgGlyValLeuThr 500
|||||
4273 TTACCTTTATGTGATGAGCGTGTCTGCTGGCCCTCCCTGACCTAACTAACCCCT 4322
501 GlnThrLeuGlyProTrpArgProValAlaTyrLeuSerLysLysLe 517
|||||
4323 CAAACTCTAGNCCATGAGGAGACCTGTGCTACCTACCTGCTCAAGAGACT 4372
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaIleV 534
|||||
4373 CGATCTCTGTAGCCAGTGGTGGCCGTATGCCGTGAAGGCTATCGCAGCTG 4422
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550
|||||
4423 TGGCCATACTGCTCAAGGAGCGTGCACAAATGACTTTGGGACAGAAATA 4472
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567
|||||
4473 ACTGTAATAGCCCCCATCGTTGGAGAACATCGTTCCGCCAGCCCCAGA 4522
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuT 584
|||||
4523 CCGATGGATGACCAACGCCCGCATGACCCACTATCAAAAGCCTGCTTCTCA 4572
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600
|||||
4573 CAGAGAGGTCACGTTCTCCACCGCGCTCTCAACCCCTGCCACTCTT 4622
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617
|||||
4623 CTGCCCTGAAGAGACTGATGAACAGTACTCATGATTGCCATCAACTATT 4672
617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634
|||||
4673 GATTGAGAGACTGGGGTCCGCAAGGACCTTACAGACATACCGCTGACTG 4722
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650
|||||
4723 GAGAAGTCTTAACCTGGTTCACTGACGGAAGCAGCTATGTGTGGAGGT 4772
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667
|||||
4773 AAGAGATGGCTGGGGCGGGTGGTGACGGGACCCCGCACGATCTGGGC 4822
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684
|||||
4823 CAGCAGCCTGCCGNAAGGAACCTTCAGCACAAAGGCTGAGCTCATGGCCC 4872
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700
|||||
4873 TCACGCAAGCTTTCGGCTGGCCGAAGGAAATCCATAAACATTTATACA 4922
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717
|||||
4923 GACAGCAGGTATGGCTTTGGCAGCTGCACACGTACACGGGGCCATCTATAA 4972
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734
|||||
4973 GCAAGGGGGTGGCTTACCTCAGCAGGAGGGAATAAAGAACAAAGAGG 5022
734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750
|||||
5023 AAATTTCTAAGCCTATTAGAAGCCTTACATTTGCCAAAAGGCTAGCTATT 5072
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767
|||||
5073 ATACACTGCTCTGGACATCAGAAAGCCCAAGATCCCATATCCAGAGGAA 5122
767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784
|||||
5123 CCAGATGGCTGACCGGCTTCCCAAGCAGCAGCCAGCGGTGTTAAACCTTC 5172
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800
|||||
5173 TGCCTATGATAGAAACACCCCAAGCCCAAGAACCCGACGACAGTACACC 5222
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817
|||||
5223 CTAGAACACTGGCAAGAGATAAAGATAGACAGTCTCTCGAGACTCC 5272
817 oGluGlyThrCystyrThrSertyrGlyLysGluIleLeuProHisLysG 834

||||| 5273 GGAAGGGACCTGCTATACCTCAGATGGGAAGAAATCCTGCCCCACAAG 5322
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850
||||| 5323 AAGGTTTGAATATATGTCACAGATACATCTCTAACCCACCTAGGAAT 5372
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867
||||| 5373 AAACACCTGCAGCAGTGGTGCAGAACATCTCTTATCATGTCTTCAGGCT 5422
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884
||||| 5423 ACCAGGAGTGGCTGATTCCGGTGGTCAAAACACTGTGTGCCCTGCCAGCTGG 5472
884 alAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900
||||| 5473 TTAATGCTTAATCTTCCAGATACCTCCAGGAAGAGACTAAGGGGAGC 5522
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917
||||| 5523 CACCCAGGCGCTCACTGGGAAGTGGACTTCACTGAGGTAAAGCGGCTAA 5572
917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934
||||| 5573 ATACGGAAACAATAATCTATTGTTGTTTGTAGACACCTTTTTCAGGATGGG 5622
934 aIGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950
||||| 5623 TAGAGGCTTATCTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAAAAA 5672
951 IleLeuGluGluIlePhePheArgPheGlyIleProLysValIleGlySe 967
||||| 5673 ATACTGGAGGAAATTTTCCAGATTGGAATACCTAAGGTAATAGGTC 5722
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984
||||| 5723 AGACAATGTCAGCTTTCGTGGCCAGGTAAAGTCAGGACTGGCCCAAGA 5772
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000
||||| 5773 TATTGGGATTCATTGGAACCTGCATTGTCATACAGACCCCAAGCTCA 5822
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017
||||| 5823 GGACAGGTAGAGAGGATGAATAGAACCATTAAGAGAGACCTTACTAAAT 5872
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034
||||| 5873 GACCGGGAGACTGGCGTTAATGATTGGATAGCTCTCCTGCCCTTTGTAC 5922
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050
||||| 5923 TTTTATAGGTTAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATGAA 5972
1051 LeuLeuTyrGlyGlyProProProLeuAlaGluIleAlaPheAlaHisSe 1067
||||| 5973 TTAATCTACGGGGGACCCCGCCCATTTGGTAGAAATTTGCTTCGATACATAG 6022
1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084
||||| 6023 TGCTGACGTGCTGCTTCCAGCCCTTGTCTCTAGGCTCAAGGCATTTG 6072
1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSer... 1099
||||| 6073 AGTGGGTGAGACAACAGCGGTGGAGCAACTCCGGGAGGCGCTTACTCAGGA 6122
1100 GlyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTy 1116
||||| 6123 GGAGAGACTTCAGATCCCATCGTTTCCAAAGTGGGAGATTCAGTCTA 6172
1116 rValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProT 1133
|||||

6173 CGTTAGACGCCACCGTGCAGAGAAACCTCGAGACTCGTGGTGAAGGCCCTT 6222
1133 yrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIlePro 1149
||||| 6223 ATCACCTACTTTTGACCAACACACGGCTGTGAAAGTCGAAGAAATCTCC 6272
1150 ThrTrpIleHisAlaSerHisValLysProAlaProProProAspSerG 1166
||||| 6273 ACCTGGATCATGCATCCACCGTTAAGCGGGCGCCACCTCCCGATTCCGG 6322
1166 yTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgL 1183
||||| 6323 GTGGAAGCGGAAAGACTGAAATCCCTTAAGCTTCGCTCCATCCGCG 6372
1183 euValProTyrSerAsnAsnSer 1191
||||| 6373 TGGTTCTTACTCTGTCAATAACTCC 6398

seq_name: /SID82/gcgdata/geneseq/geneseqn/NA2001.DAT:AAC67023

seq_documentation_block:

ID AAC67023 standard; DNA; 7873 BP.

AC AAC67023;

XX 27-MAR-2001 (first entry)

DE PERV env protein coding sequence SEQ ID NO: 23.

XX Xenotransplantation; infectious agent; vaccine; ds.

XX Porcine endogenous retrovirus.

XX WO200071726-A1.

XX 30-NOV-2000.

XX 24-MAY-2000; 2000WO-US14296.

XX 24-MAY-1999; 9905-0135631.

XX (MAYO-) MAYO MEDICAL VENTURES.

XX Federspiel MJ;

XX WPI; 2001-032041/04.

XX Inhibiting or preventing infectious agent transmission in mammalian
PT transplant recipients, by introducing recombinant DNA comprising DNA
PT encoding extracellular proteins of the agent into donor cells, such as
PT swine cells -

XX Claim 16; Page 112-115; 144pp; English.

XX The present invention provides a method to prevent the transmission of
CC infectious agents during xenotransplantation. This involves introducing
CC to donor swine cells a recombinant DNA encoding a peptide fragment from
CC the infectious agent, and then introducing these cells into the
CC transplant recipient.

XX Sequence 7873 BP; 2200 A; 1913 C; 1962 G; 1798 T; 0 other;

alignment_scores:

Quality: 5858.00 Length: 1194

Ratio: 5.103 Gaps: 4

Percent Similarity: 96.147 Percent Identity: 94.054

alignment_block:

US-09-171-553B-5 x AAC67023 ..

Align seg 1/1 to: AAC67023 from: 1 to: 7873

```
1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17
|||||
1637 GGAAGACGGGGTTCGGACCCCTCCCGAGCCCGGTAACCTTTGAAGGT 1686
|||||
17 lGluGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34
|||||
1687 GGAGGGCAACCAAGTTGAGTCTCTGGTTGATACCGGAGCGAAACATTCAG 1736
|||||
34 alLeuLeuGlnProLeuGlyLysLeuLysGlyLysSerTrpValMet 50
|||||
1737 TGCTACTACAGCCATTAGGAATACTAAAGATAAAATAATCCCTGGGTGATG 1786
|||||
51 GlyAlaThrGlyGlnArgGlnTrpTrpThrArgArgThrValAs 67
|||||
1787 GGTGCCACAGGGCAACAACAGTATCCATGGACTTACCCGAAGAACAGTTGA 1836
|||||
67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysp 84
|||||
1837 CTTGGGAGTGGGACGGGTAAACCACCTCGTTCTTGGTCATACCTGAGTGCC 1886
|||||
84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100
|||||
1887 CAGCACCCCTCTTAGTACAGACTTATCCCAAGATGGGACCAACAATT 1936
|||||
101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117
|||||
1937 TCTTTTGAACAAGGAAACCAAGGTCTGCAATAAACAAACCTATATCAC 1986
|||||
117 rValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnV 134
|||||
1987 TGTGTTGACCCCTCCAATTAGATGACGAATATCGACTATATCTCTCCCTAG 2036
|||||
134 alLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150
|||||
2037 TAAAGCCTGATCAAAATATACAAATCTGTTGGAAACAGTTTCCCAAGCC 2086
|||||
151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProGlnVa 167
|||||
2087 TGGGCAAGAACCCGACGGGATGGTTTGGCAAGCAAGTTCCCCCAACAGT 2136
|||||
167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184
|||||
2137 TATTCAACTGGAAGGCGAGTGCACACCACTGTCAGTCAGACAGTACCCT 2186
|||||
184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200
|||||
2187 TGAGTAAGAAGCTCAAGAGGAATTCGCCCAATGTCCAAAGATTAATC 2236
|||||
201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217
|||||
2237 CAACAGGGCATCCTAGTTCTCTCCCAATCTCCCTGGAATACTCCCTGCT 2286
|||||
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234
|||||
2287 ACCGGTTAGAAAGCTTGGGACTTAATGACTATGACCAGTACAGGACTTGA 2336
|||||
234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250
|||||
2337 GAGAGTCAATAAAGCGGTGAGGATATACACCCCAACAGTCCGAAACCT 2386
|||||
251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValLe 267
|||||
2387 TATAACCTCTTGTGTGCTCTCCACCCCAACGAGCTGGTATACAGTATT 2436
|||||
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnp 284
|||||
2437 GGACITTAAGGATGCTTTTCTGCTCAGATTACACCCCACTAGCCCAAC 2486
|||||
284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
|||||
2487 CACTTTTTCCTTCAATGGAGATCCAGGTACGGGAAGAACCGGCGAG 2536
|||||
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317
|||||
```

```
2537 CTCACCTGACCCGACTCCCAAGGTTCAAGAACTCCCGACCATCTT 2586
|||||
317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
|||||
2587 TGACGAAGCCCTACACAGAGACTGGCCAACTTCAGGATCCAACACCCCTC 2636
|||||
334 lInValThrLeuLeuGlnTrpValAspAspLeuLeuAlaGlyAlaThr 350
|||||
2637 AGGTGACCCCTCTCCAGTACGTGGATGACCTTCTGGCGGAGCCACCC 2686
|||||
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuGluLeuSerAs 367
|||||
2687 AAACAGGACTGCTTAGAAAGCACAGGACACTACTGCTGGAATGTCTGA 2736
|||||
367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
|||||
2737 CCTAGGCTACAGAGCTCTGTCTAAGAGGCCAGATTTGCAGGAGAGAGG 2786
|||||
384 alThrTrpLeuGlyTyrSerLeuArgGlyGlnArgTrpLeuThrGlu 400
|||||
2787 TAACATACCTTGGGGTACAGTTTGGGGGACGGGACGATGGCTGACGGAG 2836
|||||
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG 417
|||||
2837 GCACGGAAGAAACTGTAGTCCAGATACCGGCCCAACACAGCCAAACA 2886
|||||
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
|||||
2887 AGTGAGAGAGTTTTTGGGACAGCTGGATTTTGCAGACTGTGGATCCCGG 2936
|||||
434 lyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450
|||||
2937 GGTTTGCCACCTTAGCAGCCCACTCTACCCACTAACCAAGAAAGAAAGG 2986
|||||
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467
|||||
2987 GAATTCCTCGGCTCCTGACCAACAGAGCATTTGATGCTATCAAAA 3036
|||||
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484
|||||
3037 GGCCTGTGTAGCGCACCTGCTCTGGCCCTCCCTGAGCTAACTAACCCCT 3086
|||||
484 heThrLeuTyrValAspGluArgLysGlyValAlaAlaArgGlyValLeuThr 500
|||||
3087 TTACCCCTTATGTGGATGAGGTAAAGGAGTAGCCCGGGGAGTTTAAAC 3136
|||||
501 GlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLe 517
|||||
3137 CAAACTCTAGACCATGGAGGAGACCTGTTGCCCTACCTGTCAAGAAGACT 3186
|||||
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaAlay 534
|||||
3187 CGATCCTGTAGCCAGTGGTTGGCCGCTATGCTGAAGGCTATCGCAGCTG 3236
|||||
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550
|||||
3237 TGGCCATACTGGTCAAGGACGCTGACAAATTGACTTTGGGACAGAAATATA 3286
|||||
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567
|||||
3287 ACTTAATAGCCCCCATGGTGGAGAACATCGTTCCGCGAGCCCCCAGA 3336
|||||
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeu 584
|||||
3337 CCGATGATGACCAACCGCCGATGACCCACTATCAAAAGCCTGCTTCTCA 3386
|||||
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600
|||||
3387 CAGAGAGGGTACAGTTCGCTCCACCGCCGCTCTCAACCCCTGCACCTCT 3436
|||||
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617
|||||
```

3437 CTGCCTGAAGAGACTGATGATCAACCAAGTCACTCATGATTGCCATCAACTATT 3486
617 uileGluThrGlyValArgLysAspLeuThrAspLeuProLeuThrG 634
3487 GATTGAGGAGACTGGGTCCGCAAGGACCTTACAGACATACCGCTGACTG 3536
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrrValValGluGly 650
3537 GAGAAGTGTAAACCTGTTCACTCAACGGAAGCAGCTATGATGGAAGGT 3586
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667
3587 AAGAGGATGGCTGGGGCGCGGTGGTGGACGGGACCGCAGATCTCGGC 3636
667 aserSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAla 684
3637 CAGCAGCCTGCCGGAAGGAGTTTCAGACAAAAGGCTGAGCTCATGGCC 3686
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrr 700
3687 TCAGCAAGCTTTCGGCTGGCCGAAGGAAATCCATAAATTTATACA 3736
701 AspSerArgTyrrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrrLy 717
3737 GACAGCAGGTATGCTTTGGCACTGCACACGTACACGGGGCCATCTATAA 3786
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734
3787 GCAAAGGGGGTGTCTACCTCAGCAGGGAGGGAATTAAGAACAAAGAGG 3836
734 luileLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750
3837 AAATCTTAAGCCTATTAGAAGCCTTACATTTCGCAAAAAGGCTAGCTATT 3886
751 lleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767
3887 ATACACTGCTCGGACATCAGAAAGCCAAAGATCCCATATCCAGAGGGA 3936
767 nGlnMetAlaAspArgValAlaLysGlnAlaGlnAlaValAlaLeuL 784
3937 CCAGATGGCTGACCGGGTTGCCAAGCAGCAGCAGCCAGGGTGTAAACCTTC 3986
784 euProIlelleGluThrProLysAlaProGluProArgArgGluTyrrThr 800
3987 TGCCTATGATAAACAACCCCAAGCCCAAGAACGACGACGACGACATACAC 4036
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817
4037 CTAGAAGACTGCAAGAGATATAAAGATAGACACGATCTCTGAGACTCC 4086
817 oGluGlyThrCystTyrrThrSerTyrrGlyLysGluIleLeuProHisLysG 834
4087 GGAAGGGACCTGCTATACCTCAGATGGGAAGGAAATCCTGCCCAAAAG 4136
834 luGlyLeuGluTyrrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850
4137 AAGGCTTGAATATGTCACACAGATACATCTCTTAACCCACCTAGGAAC 4186
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrrHisValLeuArgLe 867
4187 AAACACCTGCACAGTGTGGTCAGAACATCTCTCTATCATGTTCTGAGGCT 4236
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeu 883
4237 ACCAGAGTGGCTGATTTCGGTGGTCAACACTGTGTGGCTGCCAGCTGG 4286
884 ValAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySe 900
4287 GTA.....AAGCGGCTA 4289
900 rHisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaL 917
4290AAGCGGCTA 4299

917 ysTyrglyAsnLysTyrrLeuLeuValPheValAspThrPheSerGlyTrp 933
4300 AATACGGAACAATAATCTATTGGTTTTGTAGACACCTTTTCAGGATGG 4349
934 ValGluAlaTyrrProThrLysLysGluThrSerThrValValAlaLysLy 950
4350 GTAGAGGCTTATCCTACTAAGAAAGAGACTTCAACCGTGGTGGCTTAGAA 4399
950 stileLeuGlu.GluilePheProArgPheGlyIleProLysValIleGly 966
4400 AATACTGGAGGAATTTTCCAAAGATTGGAAATACCTAAGGTAAATAGG 4449
967 SerAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLy 983
4450 TCAGACAATGGTCCAGCTTTCGTTGCCAGTAAGTCAAGGAGCTGCCCAA 4499
983 stileLeuGlyIleAspTrpLysLeuHisCysAlaTyrrArgProGlnSerS 1000
4500 GATATTGGGGATTTGATTGAAACTGCATTTGCATACAGACCCCAAGCT 4549
1000 erGlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLys 1016
4550 CAGGACAGGTAGAGAGGATGAATAGAACCATTAAAGAGACCTTACTAAA 4599
1017 LeuThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheVa 1033
4600 TTGACGCGGAGACTGGCGTTAATGATTGGATAGCTCTCTCCCTCCCTTGT 4649
1033 lleuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrg 1050
4650 GCTTTTAGGGTTAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATA 4699
1050 luLeuLeuTyrrGlyGlyProProLeuAlaGluIleAlaPheAlaHis 1066
4700 AATTACTTACGGGGACCCCCCATTTGGTAGAAATTTGCTTCCGTACAT 4749
1067 SerAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLe 1083
4750 AGTGCTGACGTGCTTTCCTCCAGCCTTTGTTCTTAGGCTCAAGGCACT 4799
1083 uGluTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrrSer 1099
4800 TGAGTGGTGAGACACAGCAGCGTGAGGCACTCCCGGAGGCTTACTCAG 4849
1100 ..GlyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerVal 1115
4850 GAGGAGGAGACTTGCAGATCCCATCCATCGTTTCCAAAGTGGGAGATTGATC 4899
1116 TyrValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyPr 1132
4900 TAGCTTAGACGCCACCGTGCAGGAACCTCGAGACTCGGTTGGAAGGCCCC 4949
1132 oTyrrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIleP 1149
4950 TTATCAGCTACTTTTGACCAACACCAAGCTGTGAAGTCGAAGGAATCT 4999
1149 roThrTrpIleHisAlaSerHisValLysProAlaProProAspSer 1165
5000 CCACCTGGATCCATGCATCCACGTTAAGCGCGCCACCTCCCGATTCG 5049
1166 GlyTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisAr 1182
5050 GGGTGGAAAGCCGAAAGACTGAAATCCCTTAAAGCTTAAAGCTCCATCG 5099
1182 gLeuValProTyrrSerAsnAsnSer 1191
5100 CGTGGTCTCTACTCTCTCAATAACTCC 5127

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT: AAC67032
seq_documentation_block:
ID AAC67032 standard; DNA; 6076 BP.

XX AAC67032;
 AC 27-MAR-2001 (first entry)
 DT PERV env protein coding sequence SEQ ID NO: 32.
 XX
 DE Xenotransplantation; infectious agent; vaccine; ds.
 KW Porcine endogenous retrovirus.
 XX OS
 XX WO200071726-A1.
 PN 30-NOV-2000.
 XX
 XX 24-MAY-2000; 2000WO-US14296.
 XX
 XX 24-MAY-1999; 99US-0135631.
 XX (MAYO-) MAYO MEDICAL VENTURES.
 PA
 XX Federspiel MJ;
 XX
 XX WPI; 2001-032041/04.
 DR
 XX
 XX Inhibiting or preventing infectious agent transmission in mammalian
 PT transplant recipients, by introducing recombinant DNA comprising DNA
 PT encoding extracellular proteins of the agent into donor cells, such as
 PT swine cells -
 XX
 XX Claim 16; Page 117-119; 144pp; English.
 PS
 XX The present invention provides a method to prevent the transmission of
 CC infectious agents during xenotransplantation. This involves introducing
 CC to donor swine cells a recombinant DNA encoding a peptide fragment from
 CC the infectious agent, and then introducing these cells into the
 CC transplant recipient.
 XX
 SQ Sequence 6076 BP; 1619 A; 1505 C; 1598 G; 1354 T; 0 other;

alignment_scores:
 Quality: 5646.50 Length: 1106
 Ratio: 5.157 Gaps: 1
 Percent Similarity: 99.005 Percent Identity: 97.107

alignment_block:
 US-09-171-553b-5 x AAC67032 ..
 Align seg 1/1 to: AAC67032 from: 1 to: 6076

1 GlyArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17
 2758 GGAAGACGGGGTTCGGACCCCTCCCGAGCCAGGGTAACATTGAAGT 2807
 17 lGluGlnProValGluPheLeuValAspThrGlyAlaGluHisServ 34
 2808 GGAGGGCAACCCAGTGTAGTTCTGTTGATACCGGAGCGAACAATTCAG 2857
 34 aileuLeuGlnProLeuGlyLysLeuLysGluLysLysSerTrpValMet 50
 2858 TGCTACTACAGCCATTAGAAACCTAAAGATAAAAATCCCTGGGTGATG 2907
 51 GlyAlaThrGlyGlnArgGlnProThrProThrThrArgArgThrValas 67
 2908 GGTGCCACAGGGCAACACAGTAGTCCATGGACTACCCGAGAACAGTTGA 2957
 67 pleuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysp 84
 2958 CTTGGGAGTGGGACGGGTAAACCCACTCGTTCTGTGTCATACGTAGTGCC 3007
 84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetClyAlaGlnIle 100
 3008 TAACTACTTTGGGTACAGTTTGGGGACGGGACGGGACGATGCCTGACGGAG 3957

3008 CAGCACCCCTCTAGTGAGACTTATTGACCAAGATGGGACACAAATT 3057
 101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIlePh 117
 3058 TCTTTTGAACAAGGAAACAGAAAGTGTCTGCAATAACAAACCTATCAC 3107
 117 rValLeuThrLeuGlnLeuAspAspGluThrArgLeuThrSerProGlnV 134
 3108 TGTGTTGACCCCTCAATTAGATGAGCAATATCGACTATATCTCCCTAG 3157
 134 aLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150
 3158 TAAAGCCTGATCAAAATATACAAATCTGTTGGAACAGTTTCCCAAGCC 3207
 151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnVa 167
 3208 TGGGCAGAAACCGCAGGGATGGTTTGGCAAGCAAGTTCCCCCAAGT 3257
 167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184
 3258 TATTCAACTGAAGCCAGTGCACACAGTGTCACTCAGACAGTACCCCT 3307
 184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200
 3308 TCGTAAAGAAAGCTCAAGAAGGAATTCGGCGCATGTCCAAAGATTATC 3357
 201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217
 3358 CAACAGGCATCTCTGTTCTGTCATCTCCCTGGAATCTCCCTGCT 3407
 217 uProValArgLysProGlyThrAsnAspThrArgProValGlnAspLeu 234
 3408 ACCGGTTAGAAGCCTGGGACTAATGACTATCGACCATCAGGACTTGA 3457
 234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250
 3458 GAGAGTCAATAAACGGGTGCAGGATATACACCAACAGTCCCGAACCT 3507
 251 TyrAsnLeuLeuSerAlaLeuProGluArgAsnTrpThrValLe 267
 3508 TATAACCTCTTGTGTCTCTCCACCCCAACGGAGCTGGTATACAGTATT 3557
 267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284
 3558 GGACTTAAAGGATGCCITTTTCTGCTGAGATTACACCCCACTAGCCCAAC 3607
 284 roLeuPheAlaPheGluThrPArgAspProGlyThrGlyArgThrGlyGln 300
 3608 CACTTTTTCCTTCGAATGGAGATCCAGTACGGGAAGAACCGGCGAG 3657
 301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317
 3658 CTCACCTGGACCCACTGCCCCCAAGGGTTCAAGAACTCCCGGACCATCTT 3707
 317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
 3708 TGACCAAGCCCTACACAGAGACCTGGCCAACTTCAGGATCCACACCCCTC 3757
 334 lnValThrLeuLeuGlnTrpValAspAspLeuLeuAlaGlyAlaThr 350
 3758 AGGTGACCCCTCTCCAGTACGTGGATGACCTGCTTCTGGGGGAGCCACC 3807
 351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAs 367
 3808 AAACAGACTGCTTAGAAGGCACGACGACACTACTGCTGGAATTGCTCGA 3857
 367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
 3858 CCTAGGTACAGAGCCCTCTGCTAAGAAGGCCAGATTTCAGGAGAGAGG 3907
 384 alThrTyrLeuGlyTyrSerLeuArgGlyGlnArgTrpLeuThrGlu 400
 3908 TAACTACTTTGGGTACAGTTTGGGGACGGGACGGGACGATGCCTGACGGAG 3957

401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysGI 417
|||||
3958 GCACGGAAGAAACGTGTAGTCCAGTACCGGCCCAACACAGCCAAACA 4007
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
|||||
4008 AGTGAGAGAAATTTTGGGGCAGCTGGATTTTGCAGACTGTGGATCCCGG 4057
434 lyPheAlaThrLeuAlaAlaProLeuTyrrProLeuThrLysGluLysGly 450
|||||
4058 GGTITGCGACCTTACGACCCCACTCTACCCACTAACCAAGAAAGGG 4107
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467
|||||
4108 GAATTCCTCGTGGCTCTGAGCACCAGAGGCATTTGATGCTATCAAAA 4157
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484
|||||
4158 GGCCCTGTGAGCGCACCTGTCTGGCCCTCCCTGACGTAACCTAAACCT 4207
484 heThrLeuTyrrValAspGluArgLysGlyValAlaAtqGlyValLeuThr 500
|||||
4208 TTACCCCTTATGTGATGAGCGTAAAGGAGTAGCCCGGGAGTTTAAACC 4257
501 GlnThrLeuGlyProTrpArgProValAlaIleLeuSerLysLysLe 517
|||||
4258 CAAACTCTAGGACCATGAGGAGACCTCTTGCTACCTGTCAAGAAAGCT 4307
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaLav 534
|||||
4308 CGATCCTGTAGCCAGTGTGGCCGCTATGCTGAAGGCTATCGCAGCTG 4357
534 aIaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550
|||||
4358 TGGCCATACGTGTGAAGGACGCTGACAAATGACTTTGGGACAGATATA 4407
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567
|||||
4408 ACTGTAATAGCCCCCATCGTGTGGAGAACATCGTTCGGCAGCCCCAGA 4457
567 parGtrpMetThrAsnAlaArgMetThrHisTyrrGlnSerLeuLeuLeu 584
|||||
4458 CCGATGGATGACCAACGCCCGCATGCCCACTATCAAGCCGTGCTCTCA 4507
584 hrGluArgValThrPheAlaProAlaAlaLeuAsnProAlaThrLeu 600
|||||
4508 CAGAGAGGTCACGTTCGCTCCACAGCCGCTCTCAACCCCTGCCACTCT 4557
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617
|||||
4558 CTGCCTGAAGAGACTGATGAACCCAGTGCATCATGATTGCCATCAACTAT 4607
617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634
|||||
4608 GATTGAGGAGACTGGGGTCCGCAAGACCTTACAGACATACCGCTGACTG 4657
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrrValValGluGly 650
|||||
4658 GAGAACTGTTACCTGGTTCACTGACGAGGAGCAGCTATGTGTGGAGGT 4707
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667
|||||
4708 AAGAGGATGGCTGGGGCGGGTGGTGGACGGGAAATCCATAAACATTTATA 4757
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684
|||||
4758 CAGACACCTGCCGGAAGAACTTCACACAAAGGCTGAGCTCATGGCCC 4807
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrrThr 700
|||||
4808 TCACGCAAGCTTTGGCGCTGGCGGAGGAAATCCATAAACATTTATAACA 4857

701 AspSerArgTyrrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrrLy 717
|||||
4858 GACAGCAGGTATGCTTTCGACTGCACACGTACACGGGGCCATCTATATAA 4907
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734
|||||
4908 GCAACGGGGTGTCTTACCTCAGCAGGAGGAAATAAAGAACAAAGAGG 4957
734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750
|||||
4958 AAATTCCTAAGCCCTATTAGAAGCCCTTACATTTGCCAAAAAGGCTAGCTATT 5007
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767
|||||
5008 ATACACTGTCTGGACATCAGAAAGCCAAAGATCCCATATCCAGAGGAA 5057
767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784
|||||
5058 CCAGATGGCTGACCGGTTGCCAAGCAGGAGCCAGGGTGTAACTTC 5107
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrrThr 800
|||||
5108 TGCCTATGATAGAAACACCCAAAGCCCAAGACCCGAGCAGACAGTACACC 5157
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817
|||||
5158 CTAGAGACTGGCAAGAGATAAAAAGATAGACCACTCTCTGAGACTCC 5207
817 oGluGlyThrCysTyrrThrSerTyrrGlyLysGluIleLeuProHisLysG 834
|||||
5208 GGAAGGACCTGCTATACCTCAGATGGAGGAAATCCTGCCCCACAAAG 5257
834 luGlyLeuGluTyrrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850
|||||
5258 AAGGTTTAGAATATGTCCAACAGATACATCGTCTAACCACCTTAGGAAC 5307
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrrHisValLeuArgLe 867
|||||
5308 AAACACCTGCAGCAGTTGGTGCAACATCTCCTTATCATATGTTCTGAGGCT 5357
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884
|||||
5358 ACCAGGAGTGGCTGATTGCTGTGTCAAACACTGTGTGCCCTGCCAGCTGG 5407
884 alAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900
|||||
5408 TTAATGTCTAATCTCTCCAGAAATACCTCCAGAAAGAGACTAAGGGGAAGC 5457
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917
|||||
5458 CACCCAGGCGCTCACTGGGAAGTGGACTTCACTGAGGTAAAGCCGGCTAA 5507
917 sTyrrGlyAsnLysTyrrLeuLeuValPheValAspThrPheSerGlyTrpV 934
|||||
5508 ATACGAAACAAATATCTATTGGTTTTTGTAGACACCTTTTCAGGATGGG 5557
934 alGluAlaTyrrProThrLysLysGluThrSerThrValValAlaLysLys 950
|||||
5558 TAGAGGCTTATCTTACTAAGAAAGAGACTTCAACCCGTGGTGGCTAAGAAA 5607
951 IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967
|||||
5608 ATACTGGAGGAAATTTTCCAGATTTGGAATACCTAAGGTAAATAGGGTC 5657
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984
|||||
5658 AGACAAATGGTCCAGCTTTTCGTTGCCAGGTAAGTCAAGGACATGGCCCAAG 5707
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrrArgProGlnSerSer 1000
|||||
5708 TATTGGGAAATTAATTTGGAATCTGATTTGCTATACAGACCCCAAGCTCA 5757
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017


```

|||||
5758 GGACAGCTAGAGGATCAATAGAACATTAAAGAGACCCCTTACTAAATT 5807
|||||
1017 uThrThrGluThrClyLeuAsnAspTrpMetAlaLeuLeuProPheValL 1034
|||||
5808 GACCGCGAGAGCTGGCGTTAATGATGGATAGCTCTCCCTGCGCTTTGTGC 5857
|||||
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050
|||||
5858 TTTTATGGGTAGAACACCCCTGGACAGTTTGGGCTGACCCCTATGAA 5907
|||||
1051 LeuLeuTyrGlyGlyProProLeuAlaGluLeuAlaPheAlaHis 1067
|||||
5908 TTACTCTACGGGAGACCCGCCCATTTGGTAGAAATTGCTTCCGTACATAG 5957
|||||
1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084
|||||
5958 TGCTGAGCTGCTCTTCCAGCCTTTGTTCTTAGGCTCAAGGCATTG 6007
|||||
1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSer... 1099
|||||
6008 AGTGGGTGAGAACACGAGCGTGGAGCAACTCCGGGAGGCCTACTCAGGA 6057
|||||
1100 GlyGlyAspLeuGlnVal 1105
|||||
6058 GGAGGAGACTTGCAGATC 6075

```

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT: AAT74811

seq_documentation_block:

ID AAT74811 standard: cDNA; 8060 BP.

AC AAT74811;

DT 11-FEB-1998 (first entry)

DE Porcine retrovirus Tsukuba-1 cDNA.

KW Retrovirus; porcine; GAG protein; POL protein; ENV protein;
 KW xenotransplantation; infectious; provirus; organ transplant; donor;
 KW activated virus; Tsukuba-1; PCR; ss.

OS Porcine retrovirus.

FH Key Location/Qualifiers

FT CDS 86..2002

FT /*tag= a

FT /note= "putative ENV protein"

FT CDS 3112..4686

FT /*tag= b

FT /note= "putative GAG protein"

FT CDS 4871..8060

FT /*tag= c

FT /note= "putative POL protein (partial)"

XX WO9721836-A1.

XX 19-JUN-1997.

XX 13-DEC-1996; 96WO-US19680.

XX 14-DEC-1995; 95US-0572645.

XX (GENO) GEN HOSPITAL CORP.

XX Fishman JA;

XX WPI; 1997-332804/30.

XX New nucleic acid from porcine retroviruses - used for detecting

XX PT viruses in transplant or other tissue and for assessing risk of

XX PT transmitting infection to graft recipient

XX

PS Claim 1; Fig 1; 128pp; English.

XX This sequence represents the purified porcine retroviral cDNA
 CC sequence of Tsukuba-1 and contains the putative coding regions for viral
 CC proteins GAG, POL and ENV. This sequence and PCR fragments generated
 CC from the sequence (see AAT74812-T74882) could be used to screen organs
 CC for porcine retroviruses prior to xenotransplantation. Transplantation
 CC can increase the likelihood of retroviral activation if intact and
 CC infectious proviruses are present. The porcine retroviral sequence can be
 CC used to generate probes to determine the level (e.g. copy number) of
 CC intact (i.e. potentially replicating) porcine provirus sequences in a
 CC strain of xenograft transplantation donors. It can be used to detect
 CC mutations, genetic lesions or viral recombinants and to determine the
 CC histological localisation of activated retrovirus. Using Polymerase Chain
 CC Reaction DNA Quantitation (Pdq) on blood mononuclear cells, infectivity
 CC titration and susceptibility testing can be performed. Ultimately animal
 CC donors without intact porcine retroviral sequences or with a lower copy
 CC number of viral elements could be selected.

XX SQ Sequence 8060 BP; 2233 A; 1959 C; 2012 G; 1856 T; 0 other;

alignment_scores:

Quality: 5637.50 Length: 1127

Ratio: 5.061 Gaps: 3

Percent Similarity: 98.846 Percent Identity: 96.628

alignment_block:

US-09-171-553B-5 x AAT74811

Align seg 1/1 to: AAT74811 from: 1 to: 8060

```

1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17
|||||
4687 GGGAGACGGGGTTCGGACCCCTCCCGAGCCAGGTAACCTTTGAAGGT 4736

17 lGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34
|||||
4737 GGAGGGGCAACCACTTGAGTTCCTGGTTGATACCGGACGGAACATTCAG 4786

34 alLeuLeuGlnProLeuGlyLysLeuLysGlyLysSerTrpValMet 50
|||||
4787 TGCTACTACAGCCATTAGGAAACTAAAGATAAAAAATCTCGGTGATG 4836

51 GlyAlaThrGlyGlnArgGlnTyrProTrpThrArgArgThrValAs 67
|||||
4837 GGTGC.ACAGGGCAACAACAGTATCCATGGACTACCCGAG.ACAGATTGA 4884

67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysp 84
|||||
4885 CTGGGAGTGGGACGGGTAAACCCACTCGTTTCTGGTCATACCTGAGTGCC 4934

84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100
|||||
4935 CAGCACCCCTCTTAGGTAGAGACTTATTGACCAAGATGGGAGCACAAATT 4984

101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117
|||||
4985 TCTTTTGAACAGGGAACACCAAGAGTGTCTGCAATTAACAAACCTATCAC 5034

117 rValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnV 134
|||||
5035 TGTGTTGACCCCTCCCAATTAGATGACGATATCGACTATATCTCTCCCTAG 5084

134 alLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150
|||||
5085 TAAAGCCTGATCAAAATATACAAATTCTGGTTGGAACAGTTTCCCCCAAGCC 5134

151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProGlnVa 167
|||||
5135 TGGGCAGAAACCGGAGGATGGGTTTGGCAAGCAAGTTCCCCCACAAGT 5184

167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184

```

||||| 5185 TATTCAACTGAAGGCGAGTCCACACAGTGTCAAGTCAGACAGTACCCCT 5234
184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuLe 200
||||| 5235 TCAGTAAAGAGACTCAAGAGGAATTCGGCGCATGTGCCAAGATTATC 5284
201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217
5285 CAACAGGCGATCCTAGTTCCTGCTCCATCTCCCTGGATACTCCCTGCT 5334
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234
5335 ACCGGTTAGAAAGCGCTGGGACTAATGACTATCGACACAGTACAGACTGA 5384
234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250
5385 GAGAGGTCAATAAAGCGGTGAGGATATACACCAACAGTCCCGAACCT 5434
251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValLe 267
5435 TATAACCTCTCTGTGCTCTCCACCCCAACGGAGCTGGTATACAGATT 5484
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284
5485 GGACTTAAAGGATGCTTCTCTGCTGAGATTACACCCCACTAGCCAAAC 5534
284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
5535 CACTTTTGGCTTCCGAATGGAGATCCAGGTACGGGAAGAACCGGCGAG 5584
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317
5585 CTCACCTGGACCGACTGCCCAAGGGTCAAGAACTCCCGACCATCTT 5634
317 eaSpGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
5635 TGACGAAGCCCTTACACAGAGACCTGGCCCACTTCAGGATCCAAACCCCTC 5684
334 InValThrLeuLeuGlnTrpValAspLeuLeuAlaGlyAlaThr 350
5685 AGGTGACCTCTCCAGTACGTGGATGACCTGCTTCTGGCGGAGCCACC 5734
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAs 367
5735 AAACAGGACTGCTTAGAAGGCACGAAGCACTACTGCTGGAATGCTGA 5784
367 pleuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
5785 CCTAGGCTACAGAGCTCTGCTAAGAGGCCAGATTTCAGAGAGAGAGG 5834
384 alThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400
5835 TAACATACTGGGTACAGTTTACGGGAGCGGCGGATGCTGACGGAG 5884
401 AlaArgLysLysThrValValGlnIleProAlaProThrAlaLysGI 417
5885 GCACGGAAGAAACTGTAGTCCAGATACCGGCCCAACACACAGCCAAACA 5934
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
5935 AATGAGAGAGTGTGGGACAGCTGGATTTTCAGACTGTGGATCCCGG 5984
434 lypPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450
5985 GGTTCGACCTTAGCAGCCCACTCTACCGCTAACCAAGAAAAAGGG 6034
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467
6035 GAATTCCTCGGCTCCTGAGCAGCAGAGGCAATTTGATGATATCAAAA 6084
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484

6085 GGCCCTGTGAGGCGACCTGTCTGTGCCCCCTCCCTGACGTAACTAAACCT 6134
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500
6135 TTACCCCTTATGTGATGAGCGTAGGGAGTAGCCGGGGAGTTTAAAC 6184
501 GlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLe 517
6185 CAAACCTAGGACCATGGAGAGACCTGTCCGCTACCTGTCAAGAGAGCT 6234
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaAlav 534
6235 CGATCTGTAGCCAGTGGTGGCCCATATGCCTGAAGGCTATCGCAGCTG 6284
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnI 550
6285 TGCCCATACTGGTCAAGGACGCTGACAAATGACTTTGGGACAAGAATAT 6334
550 eThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProA 567
6335 AACTGTAAAGCCCTCATGCTTGGAGAACATCGTTTCGGCAGCCCCAG 6384
567 spArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeu 583
6385 ACCGATGGATGACCAACGCCCGCATGCCACCTATCAAAGCCTGCTTCTC 6434
584 ThrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLe 600
6435 ACAGAGGGTCAAGTTCGCTCCCAACCAACCGCTCTCAACCTGCCACTCT 6484
600 uLeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuL 617
6485 TCTGCTGTAGAGACTGATGAACCACTGATCATGATTGCCATCAACTAT 6534
617 eutleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThr 633
6535 TGATTGAGGAGACTGGGTCCCGAAGACCTTACACACATACCGCTGACT 6584
634 GlyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluI 650
6585 GGAGAAGTGCTAACTGTTCACTGACGGAAGCAGCTATGTGTGGAAGG 6634
650 yLysArgMetAlaGlyAlaValValAspGlyThrArgThrIleTrpA 667
6635 TAAGAGATGGCTGGGCGCGGTGGTGGACGGGACCCGACCATCTGGG 6684
667 laSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAla 683
6685 CCAGCAGCTGCCGGGAGGAACCTTCAGCACAAAAGGCTGAGCTCATGGC 6734
684 LeuThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleYrTh 700
6735 CTCACGCAAGCTTTGGGCTGGCCGAAGGAAATCCATAACATTTATAC 6784
700 rAspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleYrL 717
6785 GGACAGCAGGTATGCTTTCGACTGCACACGCTACATGGGCGCCATCTATA 6834
717 ysGlnArgGlyLeuLeuThrSerAlaGlyArgGluLysLysAsnLysGlu 733
6835 AACAAAGGGGTTGCTTACCTCAGCAGGAGGAAATAAAGAAACAAAGAG 6884
734 GluIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaI 750
6885 GAAATCTAAGCCTATTAGAACGCGCTATTTACCAAAAAGGCTAGCTAT 6934
750 eileHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyA 767
6935 TATACACTGCTGGACATCAGAAAGCTAAAGATCTCATATCCAGAGAA 6984
767 snGlnMetAlaAspArgValAlaLysGlnAlaLeuAlaGlnAlaValAsnLeu 783
6985 ACCAGATGGCTGACCGGTTGCCAAGCAGCAGCCAGGGTGTAAACCTT 7034

784 LeuProfilLeuGluThrProLysAlaProGluProArgArgGlnTyrTh 800
|||||
7035 CTGCTATAATAGAAATGCCCAAGCCCAAGACCCAGACAGTACAC 7084
800 rLeuGluAspPrpGlnGluLeuLysLysIleAspGlnPheSerGluThrP 817
|||||
7085 CCTAGAGACTGGCAAGAGATAAAGATAGACCA.TTCTCTGAGACTC 7133
817 roGluGlyThrCysTyrThrSerTyrGlyLysGluLeuLeuProHisLys 833
|||||
7134 CGGAAGGAGCTGTATACCTCAGATGGGAAGAAATCCTGCCCAAAA 7183
834 GluGlyLeuGluTyrValGlnGln.IleHisArgLeuThrHisLeuGlyT 850
|||||
7184 GAAGGGTTAGAAATATGTCCAAAGATACATCGTCTAACCCACCTAGGAA 7233
850 hrLysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArg 866
|||||
7234 CTAAACACCTGCAGCAGTGTGTGAGAATCCCTTATCATGTTCTGAGG 7283
867 LeuProGlyValAlaAspSerValValLysHisCysValProCysGlnLe 883
|||||
7284 CTACAGGAGTGGCTGACTCGTGTCAACATTTGTGTGCCTGCCAGCT 7333
883 uValAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlyS 900
|||||
7334 GGTTAATGCTAATCTTCCAGATGCTCCAGGGAAGAGACTAAGGGGAA 7383
900 erHisProGlyValAlaHisTrpGluValAspPheThrGluValLysProAla 916
|||||
7384 GCCACCGGGCTTCACCTGGGAAGTGGACTTCACCTGAGGTAAAGCCGGCT 7433
917 LysTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTr 933
|||||
7434 AAATATGGAACAAATACCTATTGGTTTTGTAGACACCTTTTCAGGATG 7483
933 pValGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysL 950
|||||
7484 GGTAGAGGCTTATCTCTACTAAGAAAGAGACTTCAACCGTGGTAGCTAAA 7533
950 ysileLeuGluGluIlePheProArgPheGlyIleProLysValIleGly 966
|||||
7534 AAATCTGGAGAAATTTTTCAGATTTTGAATACCTAAGGTAAATAGGG 7583
967 SerAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLy 983
|||||
7584 TCAGACAATGGTCCAGCTTTTGTGCCAGGTAAAGTCAGGAGCTGGCCAA 7633
983 sIleLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerS 1000
|||||
7634 GATATTGGGATTCATTGGAACCTGCAATTTGCATACAGACCCCAAGCT 7683
1000 erGlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLys 1016
|||||
7684 CAGACAGGTAGAGAGATGAATAGAACCATTAAGAGACCCCTTACTAA 7733
1017 LeuThrThrGluThrGlyIleAsnAspTrpMetAlaLeuProPheVa 1033
|||||
7734 TTGACCGGAGACTGGCGTTATGATTTGATAGCTCTCTGCCCTTTGT 7783
1033 lLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrG 1050
|||||
7784 GCTTTTATAGGTTAGGAACACCCCTGGACAGTTTGGCTTGACCCCTATG 7833
1050 lLeuLeuTyrGlyGlyProProProLeuAlaGluIleAlaPheAlaHis 1066
|||||
7834 AATTACTTACGGGGGACCCCGCCCATTTGTTGATAATTCCTCTGTACAT 7883
1067 SerAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLe 1083
|||||
7884 AGTGCTGATGCTGCTTTTCCAGACCTTTGTCTCTAGGCTCAAGGCACT 7933

1083 uGluTrpValArgGlnAlaTrpLysGlnLeuArgGluAlaTyrSer. 1099
|||||
7934 TGAGTGGTGCAGACAAGAGCGTGGAGCAACTCCGGAGGCTACTCAG 7983
1100 ..GlyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerVal 1115
|||||
7984 GAGGAGGAGACTTGCAGATCCCATCGTTTCCAGTGGGAGATTCCAGTC 8033
1116 TyrValArgArgHisArgAlaGlyAsn 1124
|||||
8034 TAGCTTAGACGCCACCGTCGAGGAAC 8060

seq_name: /SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAF77725

seq_documentation_block:

ID AAF77725 standard; cDNA; 8060 BP.

XX AAF77725;

DT 23-MAY-2001 (first entry)

XX Tsukuba-1 cDNA.

XX Tsukuba-1; retrovirus; graft transplantation; xenotransplantation; ss.

OS Porcine retrovirus.

PN US6190861-B1.

PD 20-FEB-2001.

PF 13-DEC-1996; 96US-0766528.

PR 14-DEC-1995; 95US-0572645.

PA (GEHO) GEN HOSPITAL CORP.

PI Fishman JA;

DR WPI; 2001-256211/26.

PT Assessing risk of endogenous retroviruses in clinical practice and in xenotransplantation, comprises using probe sequences derived from swine or miniature swine retroviral genome -

XX Claim 1; Fig 1; 127pp; English.

CC The present invention relates to a method for screening a cell or tissue for the presence or expression of a retrovirus (RV), comprising contacting a target nucleic acid from the cell or tissue with a second nucleic acid from the present invention (e.g. the present sequence or a fragment thereof). The method is useful for RV detection and to assess graft transplantation risk. Screening of animals allows the elimination of donors with active replication of known viruses. Inactive proviruses can be detected and inactivated, allowing identification and elimination of potential human pathogens derived from swine in a manner not possible in the outbred human organ donor population and is important to the development of human xenotransplantation.

XX Sequence 8060 BP; 2233 A; 1959 C; 2012 G; 1856 T; 0 other;

alignment_scores:

Quality: 5637.50 Length: 1127

Ratio: 5.061 Gaps: 3

Percent Similarity: 98.846 Percent Identity: 96.628

alignment_block:

US-09-171-553B-5 x AAF77725

Align seg 1/1 to: AAF77725 from: 1 to: 8060

1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17

||||| 4687 GGGAGACGGGTTCGGAGCCCTCCCGAGCCAGGGTAACTTTGAAGGT 4736
17 lcluclyGlnProValGluPheLeuValAspThrGlyAlaGluHiserv 34
||||| 4737 GGAGGGCAACAGTTGAGTTCCCTGGTTGATACGGGAGCGAAACATTTCAG 4786
34 alLeuLeuGlnProLeuGlyLysLeuGlyLysLysSerTyrValMet 50
||||| 4787 TGCTACTACGCATTAGGAACCTAAAGATAAATAATCTCGGTGATG 4836
51 GlyAlaThrGlyGlnArgGlnTyrProThrThrArgArgThrValAs 67
||||| 4837 GGTGC ACAGGGCAACACATATCCATGGACTACCCGGAAG ACAGTTGA 4884
67 pleuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysp 84
||||| 4885 CTTGGAGTGGAGCGGTAACCCACTCGTTCTGGTCAATCCTGAGTGCC 4934
84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100
||||| 4935 CAGCACCCCTCTTAGTAGAGACTTATTGACCAAGATGGGAGCACAAATT 4984
101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117
||||| 4985 TCCTTTGAACAGAGGAAACACAGAGTGTCTGCAAAATACAAACCTATCAC 5034
117 rValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnV 134
||||| 5035 TGTGTTGACCTCCCAATTAGATGAGCAATATCGACTACTCTCCCTAG 5084
134 alLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150
||||| 5085 TAAAGCCTGATCAAAATATACAATTCGTGTTGGAACAGTTTCCCAAGCC 5134
151 TrpAlaGluThrAlaGlyWetGlyLeuAlaLysGlnValProGlnVa 167
||||| 5135 TGGGAGAAACCGCAGGAGTGGTTTGCAAGCAAGTTCCCCACAAAGT 5184
167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184
||||| 5185 TATTCACTGAAGCCAGTGCACACCAAGTGTCACTCAGACAGTACCCCT 5234
184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200
||||| 5235 TGAGTAAAGAAAGCTCAAGAAAGAAATCGCGCGCATGTCCAAAGATTAAATC 5284
201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217
||||| 5285 CAACAGGGCATCTAGTTCTCTGTCCTGCAATCTCCTGGAATATCTCCCTGCT 5334
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234
||||| 5335 ACCGGTTAGAAAGCCTGGGACTAATGACTATTCGACCACTACAGGACTTGA 5384
234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250
||||| 5385 GAGAGGTCAATAAAGCGGTGCGAGGATATACACCAACAGTCCCGAACCCCT 5434
251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTyrThrValLe 267
||||| 5435 TATAACCTCTGTGTGCTCTCCACCCCAACGGAGCTGGTATACAGTATT 5484
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284
||||| 5485 GGACTTAAAGGATGCTTCTTCTGCTGAGATTACACCCCACTAGCCCAAC 5534
284 roLeuPhealaphGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
||||| 5535 CACTTTTTCCTCGAATGGAGAGATCCAGGTACGGGAAGAACCGGGCAG 5584
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317
|||||

5585 CTCACCTGGACCCGACTGCCCAAGGGTTCAAGAACTCCCCGACCATCTT 5634
317 eaSpGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
||||| 5635 TGACGAAGCCCTACACAGAGACTGGCCAACTTCAGGATCCAAACACCTC 5684
334 lnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350
||||| 5685 AGGTGACCTCTCCAGTACGTGGTACCTGCTCTCTGGGGAGCCACC 5734
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAs 367
||||| 5735 AAACAGGACTGCTTAGAAGCAGCAAGGCACTACTCTGGAATGTCTCA 5784
367 pleuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgGluV 384
||||| 5785 CCTAGGCTACAGAGCCTCTGCTAAGAAGGCCAGATTTCAGAGAGAGG 5834
384 alThrTyrLeuGlyTyrSerLeuArgGlyGlnArgTrpLeuThrGlu 400
||||| 5835 TAAACATACTGGGGTACAGTTTACGGGAGCGGCGCATGGCTGACGGAG 5884
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG 417
||||| 5885 GCACGGAAGAAACTGTAGTCCAGATACCGGCCCAACACACAGCCCAACA 5934
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
||||| 5935 AATGAGAGAGTTTTGGGACAGCTGGATTTTGCAGACTGTGGATCCCG 5984
434 lyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450
||||| 5985 GGTGTCGACCTTAGCAGCCCACTCTACCCGCTAACCAAGAAAGAAAGG 6034
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysly 467
||||| 6035 GAATTCCTCGGGCTCTGAGCACCAGAGGCATTTGATGCTATCAAAA 6084
467 saLaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProP 484
||||| 6085 GGCCTCTGAGCGCACCTGCTGCGCCCTCCCTGACCTAACTAAACCCCT 6134
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500
||||| 6135 TTACCTTTATGTGATGAGCGTAAGGAGTAGCCCGGGAGTTTAAACC 6184
501 GlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLe 517
||||| 6185 CAACCCCTAGGACCATGGAGAAGACCTGTGCGCTACTCTGTCAAAGAAAGCT 6234
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaLav 534
||||| 6235 CGATCTCTGAGCCAGTGGTGGCCCATATGCCCTGAAGGCTATCGCAGCTG 6284
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGln AsnIle 550
||||| 6285 TGGCCATACCTGTCAGGACGCTGCACAAATGACTTTGGGACAAAGATAT 6334
550 eThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProA 567
||||| 6335 AACTGTAAATAGCCCCCATGCATTGGAGAACATCGTTCGGCAGCCCCCAG 6384
567 spArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeu 583
||||| 6385 ACCATGGATGACCAACGCGCATGACCCACTATCAAGCCTGCTTCTC 6434
584 ThrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLe 600
||||| 6435 ACAGAGAGGTGACGTTGCTTCCCAACACCGCTCTCAACCCCTGCCACTCT 6484
600 uLeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuL 617
||||| 6485 TCTGCCTGAAGAGACTGATGAACCAAGCTGACTCATGATTGCCATCAACTAT 6534

```
617 euileGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThr 633
|||||
6535 TGATTGAGGAGACTGGGTCGCGAAGACCTTACAGACATACCGCTGACT 6584
634 GlyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluG1 650
|||||
6585 GGAGAAGTGTCAACCTGTTCACTGACGGAAGCAGCTATGTGTGGAAGG 6634
650 YLysArgMetAlaGlyAlaValValAspGlyThrArgThrIleTrpA 667
|||||
6635 TAGAGAGTGGCTGGGCGCGGTGGGACGGGACCCCGCACATCTGGG 6684
667 laSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAla 683
|||||
6685 CCAGCAGCCTGCCGGGAGGAACITCAGCACAAAGGCTGAGCTCATGGCC 6734
684 LeuThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrth 700
|||||
6735 CTCACGCAAGCTTTGGCGCTGGCGGAAGGAAATCCATAAACATTTATAC 6784
700 rAspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrL 717
|||||
6785 GGCAGCAGGTATGCCCTTCGCACTGCACACGTACATGGGCGCATCTATA 6834
717 ysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGlu 733
|||||
6835 AACAAAGGGGTGCTTACCTCAGCAGGGAGGGAATAAAGAACAAAGAG 6884
734 GluIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAla1 750
|||||
6885 GAAATCTAAGCCTATTAGAAGCGGTACATTTACCAAAAAGGCTAGCTAT 6934
750 eileHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyA 767
|||||
6935 TATACACTGTCCTGGACATCAGAAAGCTAAAGATCTCATATCCAGAGAA 6984
767 snGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeu 783
|||||
6985 ACCAGATGGCTGACCGGTTGCCAAGCAGCAGCCCGAGGTTTAACCTT 7034
784 LeuProIleIleGluThrProLysAlaProGluProArgArgGlnTyrTh 800
|||||
7035 CTGCCTATATAGAAATGCTCAAAAGCCCAAGACCCAGACAGCAGTACAC 7084
800 rLeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrP 817
|||||
7085 CCTAGAAGACTGGCAAGAGATAAAAAGATAGACCA.TTCTCTGAGACTC 7133
817 roGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLys 833
|||||
7134 CGAAGGAGCTGCTATACCTCAGATGGGAAGGAATCTGCCCCACAA 7183
834 GluGlyLeuGluTyrValGlnGln.IleHisArgLeuThrHisLeuGlyT 850
|||||
7184 GAAGGGTTAGAATATGTCCAACAAGATACATGCTCAACCCACCTAGGAA 7233
850 hrLysHisLeuGlnLeuValIleGlyThrSerProTyrHisValLeuArg 866
|||||
7234 CTAAACACCTGGCAGCAGTGGTCAGAACATCCCTTATCATGTTCTGAGG 7283
867 LeuProGlyValAlaAspSerValValLysHisCysValProCysGlnLe 883
|||||
7284 CTACAGGAGTGGCTGACTCGGTGGTCAACATTTGTGCTCCCTGCCAGCT 7333
883 uValAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlys 900
|||||
7334 GGTAAATGCTAATCCTTCCAGATGCTCCAGGAAGAGACTAAGGGGAA 7383
900 erHisProGlyAlaHisTrpGluValAspPheThrGluValLysProAla 916
|||||
7384 GCCACCCAGCGCTCACTGGGAAGTGGACTTCACTGAGGTAAAGCGCGCT 7433
```

```
917 LysTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTr 933
|||||
7434 AAATATGGAAACAATAATACCTATTGGTTTGTAGACACCTTTTCAGGATG 7483
933 pValGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysL 950
|||||
7484 GGTAGAGGCTTATCCTACTAAGAAAGAGACTTCAACCGGTGGTAGTAAA 7533
950 ysIleLeuGluGluIlePheProArgPheGlyIleProLysValIleGly 966
|||||
7534 AAATACTGGAGAAAATTTTCCAAGATTGGAAATACCTAAGGTAATAGG 7583
967 SerAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLy 983
|||||
7584 TCAGACAATGGTCCAGCTTTTGTGCCCAGGTAAGTCAGGAGCTGGCCAA 7633
983 sIleLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSers 1000
|||||
7634 GATATTGGGATTGATTGGAACCTGCATTGTGCATACAGACCCCAAGCT 7683
1000 erGlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLys 1016
|||||
7684 CAGGACAGGTAGAGAGATGAATAGAACCATTTAAAGAGACCTTACTAAA 7733
1017 LeuThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheVa 1033
|||||
7734 TTGACCGCGGAGACTGGCGTTAATGATTGGATAGCTCTCCTGCCCTTGT 7783
1033 lLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrG 1050
|||||
7784 GCTTTTATAGGTTAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATG 7833
1050 luLeuLeuTyrGlyGlyProProProLeuAlaGluIleAlaPheAlaHis 1066
|||||
7834 AATTTACTACGGGGACCCCCCATTTGGTAGAATTTGCTCTGTATCAT 7883
1067 SerAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLe 1083
|||||
7884 AGTGCTGATGCTGCTTTTCCAGCCTTTGTTCTTAGGCTCAAGGCACT 7933
1083 uGluTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSer. 1099
|||||
7934 TCAGTGGGTGAGACAACAGCGCTGGAGGCAACTCCGGGAGGCTTACTCAG 7983
1100 ..GlyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerVal 1115
|||||
7984 GAGGAGGAGACTTGCAGATCCCATCGTTTCCAAGTGGGAGATTCAGTC 8033
1116 TyrValArgArgHisArgAlaGlyAsn 1124
|||||
8034 TACGTTAGACCCCGCTGAGGAAAC 8060
seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT.AAC67021
seq_documentation_block:
ID AAC67021 standard; DNA; 6076 BP.
XX
AC AAC67021;
XX
DT 27-MAR-2001 (first entry)
XX
DE PERV env protein coding sequence SEQ ID NO: 21.
XX
KW Xenotransplantation; infectious agent; vaccine; ds.
XX
OS Porcine endogenous retrovirus.
XX
PN WO200071726-A1.
XX
PD 30-NOV-2000.
XX
PF 24-MAY-2000; 2000WO-US14296.
XX
```

PR 24-MAY-1999; 99US-0135631.

PA (MAYO-) MAYO MEDICAL VENTURES.

PI Federspiel MJ;

XX WPI; 2001-032041/04.

XX Inhibiting or preventing infectious agent transmission in mammalian
PT transplant recipients, by introducing recombinant DNA comprising DNA
PT encoding extracellular proteins of the agent into donor cells, such as
PT swine cells -

XX Claim 16; Page 107-109; 144pp; English.

XX The present invention provides a method to prevent the transmission of
CC infectious agents during xenotransplantation. This involves introducing
CC to donor swine cells a recombinant DNA encoding a peptide fragment from
CC the infectious agent, and then introducing these cells into the
CC transplant recipient.

XX Sequence 6076 BP; 1613 A; 1512 C; 1609 G; 1342 T; 0 other;

alignment_scores:

Quality: 5635.50 Length: 1106
Ratio: 5.156 Gaps: 1
Percent Similarity: 98.825 Percent Identity: 97.016

alignment_block:

US-09-171-553B-5 x AAC67021 ..

Align seg 1/1 to: AAC67021 from: 1 to: 6076

1 GlyArgArgGlySerAspProGluProGluProArgValThrLeuLysVa 17
2758 GGAAGACGGGGTTCGGAGCCCTCCCGAGCCAGGGTAACCTTGAAGGT 2807
17 lclGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34
2808 GGAGGGCAACAGCTGGTGGTTCCTGGTGTATACCGGAGCGAACAATTCAG 2857
34 alLeuLeuGlnProLeuGlyLysLeuLysLysLysSerTrpValMet 50
2858 TGCTACTACGCCATTAGGAAACATAAAGATAAATAAATCCTGGGTGATG 2907
51 GlyAlaThrGlyGlnArgGlnTyProTrpThrArgArgThrValAs 67
2908 GTGCCACAGGGCAACAACAGATATCCATGGACTACCCGAGAACAGTTGA 2957
67 pLeuGlyValGlyArgValThrHisSerPheLeuValileProGluCysp 84
2958 CTGGGAGTGGGACGGGTAAACCCACTCGTTTCTGTATACCTGAGTGCC 3007
84 roValProLeuLeuGlyArgAspLeuThrLysMetGlyAlaGlnIle 100
3008 CAGCACCCCTCTTAGTAGAGACTTATTGACCAAGATGGGAGCACAAAT 3057
101 SerPheGlnGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117
3058 TCTTTTGAACAGGGAACCAAGAGTGTCTGTCRAATAACAACCTATAC 3107
117 rValLeuThrLeuGlnLeuAspAspGluTyArgLeuTySerProGlnV 134
3108 TGTGTTGACCTCCCAATTAGATGACGAATATCGACTATACCTCCCTAG 3157
134 alLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150
3158 TAAAGCCTGATCAAAATATACAAATCTCTGGTTGGAAACAGTTTCCCAAGCC 3207
151 TrpAlaGlnThrAlaGlyMetGlyLeuAlaLysGlnValProGlnVa 167
3208 TGGGCAGAAACCGCAGGATGGGTTTGGCAAGCAAGTTCCCCACCAAGT 3257

167 ltleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyProL 184
3258 TATTCAACTGAAGGCGAGTGCACACACAGTGTCAAGTCAAGTACCCCT 3307
184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuile 200
3308 TAGATAAAGAAAGCTCAAGAAAGAAATTCGGCGCATGTCCAAAGATTAATC 3357
201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeu 217
3358 CACAGGGCATCTAGTCTCTGTCCTCAATCTCCCTGGAATACTCCCTGCT 3407
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234
3408 ACCGGTTAGAAAGCCTGGGACTAATGACTATGACGAGTACAGGACTTGA 3457
234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250
3458 GAGAGTCAATAAAGCGGTGCAGGATATACACCAACAGTCCCGAACCT 3507
251 TyrAsnLeuLeuSerAlaLeuProGluArgAsnTrpThrValLe 267
3508 TATAACCTCTTGTGCTCTCCACCCCAACGAGCTGTGTATACAGTATT 3557
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284
3558 GGACTTAAAGGATGCCCTTTTCTGCTGAGATTACACCCACTAGCCAAAC 3607
284 roleuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
3608 CACTTTTTCCTTCGAATGAGATCCAGGATCCAGGACGGAAGAACCCGGCAG 3657
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317
3658 CTCACCTGGACCGGACTGCCCAAGGGTTCAAGAACTCCCGGACCATCTT 3707
317 eAspGluAlaLeuHisArgAspLeuAlaasnPheArgIleGlnHisProG 334
3708 TGACGAGGCCCTACACAGAGACTTGCCCAACTTCAGGATCCCAACACCTC 3757
334 lnValThrLeuLeuGlnTyValAspAspLeuLeuAlaGlyAlaThr 350
3758 AGTGACCTCTCTCCAGTACGTGGATGACCTGCTTCTGGGGGAGCCACC 3807
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAs 367
3808 AACAGGACTGCTTAGAAGGCACGAGGACACTACTGCTGGAATTTGTCTCA 3857
367 pLeuGlyTyArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
3858 CTTAGGCTACAGAGCCTCTGTAAGAAAGGCCAGATTTCGAGAGAGAGG 3907
384 alThrTyLeuGlyTySerLeuArgGlyGlnArgTrpLeuThrGlu 400
3908 TAACATACTTGGGTACAGTTTCGGGAGGGGAGGAGCATGGCTGAGCGAG 3957
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG 417
3958 GCACGGAAGAAACCTGTAGTCCAGATACCGGCCCAACACAGCCCAACA 4007
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
4008 AGTGAGAGAGATTTTGGGACAGCTGGATTTTGCAGACTGTGGATCCCGG 4057
434 lyPheAlaThrLeuAlaAlaProLeuTyProLeuThrLysGluLysGly 450
4058 GGTITGGACCTTAGCAGCCCTCTTACCCACTAACCAAGAAAGAGG 4107
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467
4108 GAATTCCTCTGGGCTCTCTGAGCAGCAGGAGGCAATTTGTGCTATCAAAA 4157

467 sAlaLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProP 484
|||||
4158 GCCCTGCTGAGCGACCTGCTCGGCCCTCCCTGACGTAACTAAACCC 4207
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500
|||||
4208 TTACCCCTTTATGTGATGAGCGTAAGGAGTAGCCCGGGAGTTTTAAC 4257
501 GlnThrLeuGlyProTrpArgProValAlaTyrLeuSerLysLysLe 517
4258 CAAACTCTAGGACCATGGAGAGACCTGTTGCCCTACCTGTCAAAGAAGCT 4307
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaAlaAlav 534
4308 CGATCCCTGTAGCCAGTGGTTGGCCGCTATGCCCTCAAGGCTATCGCAGCTG 4357
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550
4358 TGCCATACTGCTCAAGGACGCTGACAAATTGACTTTGGGACAGATATA 4407
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProAs 567
4408 ACTGTATAGCCACCATGGCTTGGAGAACATCGTTCCGGCAGCCCCAGA 4457
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuT 584
4458 CCGATGATGACCAACGCCCGCATGACCCACTATCAAAAGCCTCTCTCA 4507
584 hrGluArgValThrPheAlaProAlaAlaLeuAsnProAlaThrLeu 600
4508 CAGAGAGGTCACGTTCCGCTCCACGACGCCCTCAACCTGCCACTCTT 4557
601 LeuProGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617
4558 CTCGCTGAGAGACTGATGAACCACTGACTCATGATTGCCATCAACTATT 4607
617 uIleGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634
4608 GATTGAGGAGACTGGGCTCCGAAGGACCTTACAGACATACCGCTGACTG 4657
634 lYgluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650
4658 GAGAACTGTTAACTCTGTTCACTGACGGAAGCAGCTATGTTGGTGAAGT 4707
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667
4708 AAGAGGATGGCTGGGGCGCGGTGTGTGACGGGACCCGACCATCTGGGC 4757
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAla 684
4758 CAGCAGCTGCCGGAAGGAACCTTACGACAAAGGCTGAGCTCATGGGCC 4807
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700
4808 TCACGCAAGCTTTGGCGCTGGCGGAAGGAAATCCATAAACATTATACA 4857
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrI 717
4858 GACAGCAGGTATGCCCTTTCGCGACTGCACACGTCACACGGGGCCATCTATAA 4907
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734
4908 GCNAAAGGGGTGCTTACCCTACGAGGGGGAATAAAGAACAAAGAGG 4957
734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750
4958 AAATCTAAGCCTATTAGAAGCCTTACATTTCCCAAAAAGGCTACTATT 5007
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767
5008 ATACACTGTCTCGGACATCAGAAAGCAAAAGATCCCATATCCAGAGGAA 5057
767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeu 784

5058 CCAGATGGCTGACCGGTTGCCAAAGCAGGCCAGGGGTGTAACCTTC 5107
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800
|||||
5108 TGCTATGATAGAAACACCAAGCCCAAGCCCGGACGAGCAGTACAC 5157
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817
|||||
5158 CTAGAAGACTGGCAAGAGATAAAAGATAGACCAGTCTCTGAGACTCC 5207
817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834
5208 GGAAGGAGCTGCTATACCTCAGATGGGAAGGAAATCTGCCCCACAAG 5257
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850
5258 AAGGGTTAGAATATGTCCCAACAGATACATCGTCTAACCCACCTAGGA 5307
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867
5308 AAACACCTGCACGAGTTGCTCAGAACATCTCCTTATCATGTCTGAGCT 5357
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeu 884
5358 ACCAGAGTGGCTGATTGGTGGTCAAACTGTGTGCCCTGCCAGCTGG 5407
884 alAsnAlaAsnProSerArgIleProGlyLysArgLeuArgGlySer 900
5408 TTAATGCTAAATCCTCCAGAAATACCTCCAGAAAGAGACTAAGGGGAAG 5457
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917
5458 CACCAGGCGCTCACTGGAAAGTGACTTCACTGAGGTAAGCCGCTAA 5507
917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934
5508 ATACGGAACAAATATCTATTGGTTTTGTAGACACCTTTTCAGGATGG 5557
934 alGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950
5558 TAGAGCTTATCCTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAGAAA 5607
951 IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967
5608 ATACTGGAGGAAATTTTCCAGATTTGGAAATACCTAAGGTAAATAGGTC 5657
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984
5658 AGACAATGGTCCAGCTTTGCTGCCCAGGTAAAGTCAGGAGCTGGCCAAGA 5707
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000
5708 TATTGGGATTTGTTGAAACTGCTATGTGCATACAGACCCCAAGCTCA 5757
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017
5758 GGACAGGTAGAGAGGATGAATAGAACCATTAAAGAGACCCTTACTAAT 5807
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheVal 1034
5808 GACCGCGAGACTGGCGTTAATGATTGATAGCTCTCTCTGCCCTTTGTGC 5857
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050
5858 TTTTTAGGTTAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATGAA 5907
1051 LeuLeuTyrGlyGlyProProLeuAlaGluIleAlaPheAlaHisSe 1067
5908 TTACTCTACGGGGACCCCCCATTTGGTGAATTTGCTTCCGTACATAG 5957
1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084

US-09-171-553B-5 x AAV09698
Align seg 1/1 to: AAV09698 from: 1 to: 3320
271 AspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAl 287
22 GATGCCCTTCTTCTGCTGAGATTACACCCCACTAGCCCAACCACTTTTGC 71
287 aPheGluTrpArgAspProGlyThrGlyArgThrGlyGlnLeuThrTrpT 304
72 CTTTGAATGGAGAGATCCAGGTACGGGAAGAACCGGGCAGCTCACCTGGA 121
304 hrArgLeuProGlnGlyPheLysAsnSerProThrIlePheAspGluAla 320
122 CCGGACTGCCCAAGGGTTCAAGAACTCCCGACCATCTTTGACGAAGCC 171
321 LeuHisArgAspLeuAlaAsnPheArgIleGlnHisProGlnValThrLe 337
172 CTACACAGGACCTGCCCAACTTCAGGATCCACACCCCTCAGGTGACCT 221
337 uLeuGlnTrpValAspAspLeuLeuAlaGlyAlaThrLysGlnAspC 354
222 CCTCCAGTACGTGGATGACCTGCTTCTGCGGGAGCCACCAACAGGACT 271
354 yLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAspLeuGlyTyr 370
272 GCTTAGAAGGTACGAAGGCCTACTGCTGGAATGTCTGACCTAGGCTAG 321
371 ArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluValThrTyrLe 387
322 AGAGCCTCTGCTAAGAAGCCCAAGATTTCAGGAGAGAGGTAACATATT 371
387 uGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGluAlaArgLysL 404
372 GGGGTACAGTTTGGGGGGCGGCGAGGATGGCTGACGGAGGACGGAAGA 421
404 yThrValValGlnIleProAlaProThrThrAlaLysGlnValArgGlu 420
422 AAAGCTAGTCCAGATACCGGCCCAACACACAGCCAAACAGAGGAGAG 471
421 PheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProGlyPheAlaTh 437
472 TTTTGGGGACAGCTGGAATTTTCAGACTGTGGATCCCGGGGTTTGCAG 521
437 rLeuAlaAlaProLeuTyrProLeuThrLysGluLysGlyGlyPheSerT 454
522 CTTAGCAGCCCACTCTACCCGCTAACCAAGAAAGGGGGATTCCTCT 571
454 rAlaProGluHisGlnLysAlaPheAspAlaIleLysLysAlaLeuLeu 470
572 GGGCTCTGAGCAGCAGAGGCAATTTGATGCTATCAAAAGGCCCTGCTG 621
471 SerAlaProAlaLeuAlaLeuProAspValThrLysProPheThrLeuTy 487
622 AGCGACCTGCTCTGGCCCTCCCTGAGTAACTAAACCCCTTTACCCCTTA 671
487 rValAspGluArgLysGlyValAlaAlaArgGlyValLeuThrGlnThrLeuG 504
672 TGTGGATGAGCGTAAGGAGTAGCCCGAGGAGTTTAAACCCCAACCCCTAG 721
504 lProTrpArgArgProValAlaTyrLeuSerLysLysLeuAspProVal 520
722 GACCATGGAGGAGACCTGTTGCCCTACCTGTCAAAGAGCTTGATCCTGTA 771
521 AlaSerGlyTrpProValCysLeuLysAlaIleAlaAlaValAlaIleLe 537
772 GCCAGTGGTTGGCCCGGTATCTGAAGGCTATCGCAGCTGTGGCCATACT 821
537 uValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIleThrValIleA 554
822 GGTCAAGGAGCGCTGACAAATTTGACTTTGGGACAGATATACCTGTAATAG 871
554 laProHisAlaLeuGluAsnIleValArgGlnProProAspArgTrpMet 570

5958 TGCTGACCTGCTGCTTTCCAGCCTTTGCTCTAGGCTCAAGGCACTTG 6007
1084 luTrpValArgGlnAlaTrpLysGlnLeuArgGluAlaTyrSer... 1099
6008 AGTGGGTGAGACACAGGCGTGGAGGCACTCCGGAGGCTACTCAGGA 6057
1100 GlyGlyAspLeuGlnVal 1105
6058 GGAGGAGACTTGCATC 6075
seq_name: /SDS2/gcdata/geneseq/geneseq/NA1997.DAT.AAV09698
seq_documentation_block:
ID AAV09698 standard; DNA; 3320 BP.
XX
AC AAV09698;
XX
DT 19-MAY-1998 (first entry)
XX
DE Porcine retrovirus pol and env DNA.
XX
KW Porcine retrovirus; PoEV; POL protein; ENV protein; vaccine;
KW diagnosis; xenotransplantation; prophylactic; therapeutic; ds.
XX
OS Porcine retrovirus.
XX
FH Key Location/Qualifiers
FT CDS 23..2793
FT /tag= a
FT /product= POL protein
FT /note= "polymerase protein"
FT CDS 2642..3297
FT /tag= b
FT /product= ENV protein
FT /note= "envelope protein"
XX
XX W09740167-A1.
XX
XX 30-OCT-1997.
XX
XX 18-APR-1997; 97WO-GB01087.
XX
XX 10-FEB-1997; 97GB-0002668.
XX 19-APR-1996; 96GB-0008164.
XX
XX (IMUT-) IMUTRAN LTD.
XX (QONE-) Q-ONE BIOTECH LTD.
XX
XX Galbraith DN, Haworth C, Lees GM, Smith KT;
XX WPI; 1997-535851/49.
XX
XX Polynucleotide encoding porcine retrovirus expression product -
XX useful to develop products for use in vaccines, diagnosis and
XX xeno-transplantation
XX
XX Claim 4; Fig 1; 69pp; English.
XX
XX This DNA sequence encodes the porcine retrovirus (PoEV) polymerase (POL)
XX and envelope (ENV) proteins. These proteins can be used to develop viral
XX vaccines, antisense nucleic acids, ribozymes and other antiviral agents.
XX They can also be used in xeno-transplantation technology and as
XX diagnostic tools.
XX
XX Sequence 3320 BP; 922 A; 844 C; 832 G; 722 T; 0 other;
XX
alignment_scores:
Quality: 4840.00 Length: 921
Ratio: 5.261 Gaps: 0
Percent Similarity: 99.891 Percent Identity: 99.674
alignment_block:

|||||
872 CCCCCCATGATTCGAGAACATCTTCGGCAGCCCCCAGACCGATGGATG 921
571 ThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuThrGluArgVa 587
922 ACCAACCCCGCATGACCCACTATCAAAAGCCTGCTTCTCAGAGAGGGT 971
587 lThrPheAlaProProAlaAlaLeuAsnProAlaThrLeuLeuProGluG 604
972 CACTTTCGCTCCACCAGCCGCTCTCAACCCCTGCCACTCTCTCCCTGAAG 1021
604 luthrAspGluProValThrHisAspCysHisGlnLeuLeuLeuGluGlu 620
1022 AGACTGATGAACAGTGACTCATCATGTCATCAACTATTGATTGAGGAG 1071
621 ThrGlyValArgLysAspLeuThrAspIleProLeuThrGlyGluValLe 637
1072 ACTGGGTCGCCAAGGACCTTACAGACATACCCGCTGACTGGAGAAGTGCT 1121
637 uThrTrpPheThrAspGlySerSerTyrValValGluGlyLysArgMeta 654
1122 AACCTGGTTCACCTGACGGAAGCAGCTATGTGTGGTGAAGTAAAGAGATGG 1171
654 laglyAlaAlaValValAspGlyThrArgThrIleThrPalaSerSerLeu 670
1172 CTGGGGCGGCGAGTGGTGGAGCGGACCGCAGCATCTGGGCCAGCAGCTG 1221
671 ProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaLeuThrGlnAl 687
1222 CCGGAAGGAACCTCAGCGCAAAAGGCTGAGCTCATGGCCCTCAGCGAAGC 1271
687 aLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThrAspSerArgT 704
1272 TTTTCGGCTGGCCGAAGGAAATCCATAAACATTTATACGGACAGAGT 1321
704 yrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLysGlnArgGly 720
1322 ATGCCCTTGGAGCTGCACACGTACACGGGGCCATCTATATAACAAAGGGGG 1371
721 LeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluGluLeuLeuSe 737
1372 TTGCTTACCTCAGCAGGGAGGAAATTAAGACAAAGAGAAATTCCTAAG 1421
737 rLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIleHisCysP 754
1422 CCTATTAGAGGCTTACATTTGCCAAAAGGCTAGCTATTATACACTGTC 1471
754 roglyHisGlnLysAlaLysAspLeuIleSerArgGlyAsnGlnMetAla 770
1472 CTGGCATCAGAAAGCCAAAGATCTCATATCTAGAGGGAACAGATGGCT 1521
771 AspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuLeuProIleI 787
1522 GACCGGGTTCGAAGCAGCAGCCAGCTGTATTACCTTCTGCCTATAAT 1571
787 eGluThrProLysAlaProGluProArgArgGlnTyrThrLeuGluAspT 804
1572 AGAAAGCCCAAGGCCCAAGACCCAGACGACAGTACACCTTAGAAGACT 1621
804 rpGlnGluIleLysLysIleAspGlnPheSerGluThrProGluGlyThr 820
1622 GGCAGAGATAAAAGATAGACCACTTCTCTGAGACTCCGGAGGGGACC 1671
821 CysTyrThrSerTyrGlyLysGluIleLeuProHisLysGluGlyLeuGl 837
1672 TGCTATACCTCATATGGGAGGAATCTGCCCCCAAGAAAGAGGTTAGA 1721
837 uTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThrLysHisLeuG 854
1722 ATATGTCCAACAGATACATCGTCTAACCCACCTAGGAACCTAACACCTGC 1771
854 lnglnLeuValArgThrSerProTyrHisValLeuArgLeuProGlyVal 870
|||||
1772 AGCAGTTGGTCAGAACATCCCCCTTATCATGTTCTGAGGCTACCAGAGTG 1821
871 AlaAspSerValValLysHisCysValProCysGlnLeuValAsnAlaAs 887
1822 GCTGACTCGGTGCTCAACATTTGTGTGCCCTGCCAGCTGGTTAATGCTAA 1871
887 nProSerArgIleProProGlyLysArgLeuArgGlySerHisProGlyA 904
1872 TCTTCCAGAAATACCTCCAGGAAGAGACTAAGGGGAAGCCACCAGGCG 1921
904 laHisTrpGluValAspPheThrGluValLysProAlaLysTyrGlyAsn 920
1922 CTCACCTGGGAAGTGGACTTCACTGAGGTAAAGCCGCTAAATACGGAAC 1971
921 LysTyrLeuLeuValPheValAspThrPheSerGlyTrpValGluAlaTy 937
1972 AAATATCTATTGTTTGTAGACACCTTTTCAGGATGGTAGAGCTTA 2021
937 rProThrLysLysGluThrSerThrValValAlaLysLysLysIleLeuGluG 954
2022 TCTTACTAAGAAAGAGACTTCAACCGTGGTGAAGAAATACTGGAGG 2071
954 luIlePheProArgPheGlyIleProLysValIleGlySerAspAsnGly 970
2072 AAATTTTCCAAAGATTTGGAATACCTAAGGTAATAGGCTCAGACAATGGT 2121
971 ProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysIleLeuGlyI 987
2122 CCAGCTTTCGTTGCCAGGTAAAGTCAGGAGCTGGCCAAAGATATTGGGAT 2171
987 eaSrpLysLysLeuHisCysAlaTyrArgProGlnSerSerGlyGlnValG 1004
2172 TCAATTGGAACTGCATTGTGCATACAGACCCCAAGCTCAGACAGGTAG 2221
1004 luArgMetAsnArgThrIleLysGluThrLeuThrLysLeuThrThrGlu 1020
2222 AGAGGATGAATAGAACCATTAAGAGAGACCTTACCAAAATTGACCACAGAG 2271
1021 ThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValLeuPheArgVa 1037
2272 ACTGCATTAATGATTTGATGGCTCTCCTGCCCTTTGTGCTTTTAGGGT 2321
1037 lArgAsnThrProGlyGlnPheGlyLeuThrProTyrGluLeuLeuTyrG 1054
2322 GAGGAACACCCCTGGACAGATTGGGCTACCCCTATGAATTGCTCTACG 2371
1054 lyGlyProProProLeuAlaGluIleAlaPheAlaHisSerAlaAspVal 1070
2372 GGGGACCCCCCGTTGGCAGAAATTTGCCTTTGCACATAGTGTGATGTG 2421
1071 LeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuGluTrpValAr 1087
2422 CTGCTTTCCAGCCTTTGTTCTTAGGCTCAAGGCGCTCGAGTGGGTGAG 2471
1087 gGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSerGlyGlyAspLeuG 1104
2472 GCAGGAGCGGTGAAGCAGCTCCGGAGGCCCTACTCAGGAGAGACTTGC 2521
1104 lnValProHisArgPheGlnValGlyAspSerValTyrValArgHis 1120
2522 AAGTTCACATCGCTTCCAGTTGGAGATTGAGTCTATGTTAGACGCCAC 2571
1121 ArgAlaGlyAsnLeuGluThrArgTrpLysGlyProTyrLeuValLeuLe 1137
2572 GTGCAGGAACCTCGAGACTCGGTGGAAGGACCTTATCTCGTACTTTT 2621
1137 uThrThrProThrAlaValLysValGluGlyIleProThrTrpIleHisA 1154
2622 GACCACACCAAGCTGTGAAGTGAAGGAATCCCCACCTGGATCCATG 2671
1154 laSerHisValLysProAlaProProProAspSerGlyTrpArgAlaGlu 1170
2672 CATCCACGTTAAGCCGCGGCCACCTCCCGATTCCGGGTGGAAGCCGAA 2721

```
1171 LysThrGluAsnProLeuLysLeuArgLeuHisArgLeuValProTyrSe 1187
|||||
2722 AAGACTGAAATCCCTTACCTTCGCTCCATCGGTGCTTACTC 2771
|||||
1187 rAsnAsnSer 1191
|||||
2772 TGTCATAACTCC 2784

seq_name: /SDS2/gcgdata/geneseq/geneseq/NA2001.DAT.AAF77726
seq_documentation_block:
ID AAF77726 standard; DNA; 7333 BP.
XX
AC AAF77726;
XX
DT 23-MAY-2001 (first entry)
XX
DE Defective retroviral genome isolated from PK-15 cell line.
XX
KW Retrovirus; graft transplantation; xenotransplantation; PK-15 cell line;
XX
OS ss.
XX
PN Unidentified.
XX
PD US6190861-B1.
XX
PF 20-FEB-2001.
XX
PR 13-DEC-1996; 9605-0766528.
XX
PS 14-DEC-1995; 9505-0572645.
XX
PA (GEO ) GEN HOSPITAL CORP.
XX
PI Fishman JA;
XX
DR WPI; 2001-256211/26.
XX
P-PSDB; AAB73282, AAB73283, AAB73284.
XX
PT Assessing risk of endogenous retroviruses in clinical practice and in
PT xenotransplantation, comprises using probe sequences derived from swine
PT or miniature swine retroviral genome -
XX
PS Claim 1; Fig 2; 127pp; English.
XX
CC The present invention relates to a method for screening a cell or tissue
CC for the presence or expression of a retrovirus (RV), comprising
CC contacting a target nucleic acid from the cell or tissue with a second
CC nucleic acid from the present invention (e.g. the present sequence or a
CC fragment thereof). The method is useful for RV detection and to assess
CC graft transplantation risk. Screening of animals allows the elimination
CC of donors with active replication of known viruses. Inactive proviruses
CC can be detected and inactivated, allowing identification and elimination
CC of potential human pathogens derived from swine in a manner not possible
CC in the outbred human organ donor population and is important to the
CC development of human xenotransplantation.
XX
SQ Sequence 7333 BP; 1984 A; 1788 C; 1885 G; 1676 T; 0 other;

alignment_scores:
Quality: 4489.50 Length: 1191
Ratio: 5.033 Gaps: 2
Percent Similarity: 74.895 Percent Identity: 74.643

alignment_block:
US-09-171-553B-5 x AAF77726 ..
Align seg 1/1 to: AAF77726 from: 1 to: 7333

1 GlyArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17
|||||
```

```
2173 GGGAGAGCGGGTTCGAGACCCCTCCCGAGCCAGGGTAACCTTTGAAGGT 2222
17 IGLuGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34
|||||
2223 GGAGGGGCAACCAAGTTGAGTTCCTGGTTGATACCGGAGCGGAGCATTCAG 2272
34 alLeuLeuGlnProLeuGlyLysLeuLysGluLysLysSerTrpValMet 50
|||||
2273 TGCTGCTACAACCATTTAGGAAACTAAAAAGAAAAAATCCTGGGTGATG 2322
51 GlyAlaThrGlyGlnArgGlnTyrProTyrThrArgArgThrValAs 67
|||||
2323 GGTGCCACAGGGCAACGGCAGATATCCATGGACTACCCGAGAACCGTTGA 2372
67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysP 84
|||||
2373 CTTGGGAGTGGGACGGGTAAACCCACTCGTTTCTGTGTCATCCTCGAGTGC 2422
84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100
|||||
2423 CAGTACCCCTTCTAGTAGAGACTTACTGACCAAGATGGGAGCTCAAT 2472
101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117
|||||
2473 TCCTTTGACACAGAGAGACCAAGAGTGTCTGTGAATAACAAACCCATCAC 2522
117 rValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnV 134
|||||
2523 TGTGTTGACCCCTCCAATTTAGATGATGAATATCGACTATATCTCCCAAG 2572
134 allysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150
|||||
2573 TAAAGCCTGATCAAGATATACAGTCTCTGGTTGGAGCAGTTTCCCAAGCC 2622
151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnVa 167
|||||
2623 TGGGCAGAAACCCGAGGATGGGTTTGGCAACGCAAGTCCCCACAGGT 2672
167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184
|||||
2673 TATTCAACTGAAGGCCAGTCTACACAGTATCAGTCAGACAGATACCCCT 2722
184 euSerArgGluAlaArgGluGlyIleTyrProHisValGlnArgLeuIle 200
|||||
2723 TGAGTAGAGAGGCTCGAGAGGAATTTGGCCGATGTTCAAGATTAAATC 2772
201 GlnGlnGlyIleLeuValProValGlnSerProTyrAsnThrProLeuLe 217
|||||
2773 CAACAGGGCATCCTAGTTCCTGTCCAATCCCTTGGAAATACCTCCCTGCT 2822
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234
|||||
2823 ACCGGTTAGGAAGCCTGGGACCAATGATTATCCAGCAGTACAGGACTTGA 2872
234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250
|||||
2873 GAGAGGTCAATAAAGGTCGAGAGACATACACCAACGGTCCCGAACCCCT 2922
251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValle 267
|||||
2923 TATAACCTCTTGAGCGCCCTCCCGCCTGAACGGAACTGGTACACAGTATT 2972
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284
|||||
2973 GGACTTAAAGATGCCCTTCTTCTCCCTGAGATTACACCCCACTAGCAAC 3022
284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
|||||
3023 CACTTTTACCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCGGCAG 3072
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317
|||||
3073 CTCACCTGGACCCGACTGCCCAAGGGTTCAAGAACTCCCGCGACCATCTT 3122
```

317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
 |||||
 3123 TGAGAAAGCCCTACACAGGACCTGGCCAACTTCAGGATCCAAACACCTC 3172
 334 InValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350
 |||||
 3173 AGGTGACCTCTCCAGTACGTGATGACCTGCTTCTGGGGAGCCACC 3222
 351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuLeuSerAs 367
 |||||
 3223 AAACAGACTGCTTAGAAGGTACGAAGGCACCTACTGCTGGAATTGCTGA 3272
 367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
 |||||
 3273 CCTAGGCTACAGAGCCTCTCTAAGAAGGCCAGATTTCAGAGAGAGAGG 3322
 384 alThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400
 |||||
 3323 TAACATACCTTGGGGTACAGTTTGGGGGGGGGAGCATGGCTGACGGAG 3372
 401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysGI 417
 |||||
 3373 GCAGGNAGAAACTGTAGTCCAGATACCGCCCCCAACACACAGCCAAACA 3422
 417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
 |||||
 3423 AGTGAGAGAGCTTTTGGGACACCTGGATTTTGCACACTGTGGATCCCGG 3472
 434 lyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450
 |||||
 3473 GGTTCGACCTTAGCAGCCCCACTCTACCGGCTAACCAAGAAAAAGGG 3522
 451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysly 467
 |||||
 3523 GGTG..... 3527
 467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484
 3527 3527
 484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500
 3527 3527
 501 GlnThrLeuGlyProTrpArgProValAlaTyrLeuSerLysLysIle 517
 3527 3527
 517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaAlav 534
 3527 3527
 534 aAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550
 3527 3527
 551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567
 3527 3527
 567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeu 584
 3527 3527
 584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600
 3527 3527
 601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617
 3527 3527

617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634
 3527 3527
 634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650
 3527 3527
 651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667
 3527 3527
 667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684
 3527 3527
 684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700
 3527 3527
 701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717
 3527 3527
 717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734
 |||||
 3528CTTACTCTCAGCAGGGAGGGAATAAAGACAAGAGG 3564
 734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750
 |||||
 3565 AAATTCCTAAGCCTATTAGAAGCCTTACATTTGCCAAAAAGGCTAGCTATT 3614
 751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767
 |||||
 3615 ATACACTGCTCTGGACATCAGAAAGCCAAAGATCTCATATCTAGAGGAA 3664
 767 nglnMetAlaAspArgValAlaLysGlnAlaGlnAlaValAsnLeuL 784
 |||||
 3665 CCAGATGGCTGACCGGTTGCCAAGCAGCAGCCCGAGCTGTTAACCTTC 3714
 784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800
 |||||
 3715 TGCTCTATAAGAAAGCCCAAGCCCAAGCCAGACAGACAGTACACC 3764
 801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817
 |||||
 3765 CTAGAAGACTGGCAAGAGATAAAAAAGATAGACCAGTTCTCTGAGACTCC 3814
 817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834
 |||||
 3815 GGAGGGGACCTGCTATACCTCATATGGGAAGAAATCTGCCCCACAAG 3864
 834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850
 |||||
 3865 AAGGTTAGAATATGTCCACAGATACATCGCTAACCCACCTAGGAAC 3914
 851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867
 |||||
 3915 AAACACCTGCAGCAGTTGGTCAGAACATCCCTTATCATGTCTGAGGCT 3964
 867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884
 |||||
 3965 ACCAGGAGTGGCTGACTCGGTGGTCAACAATGTGTCCTGCCAGCTGG 4014
 884 alAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900
 |||||
 4015 TTAATGCTAATCCTTCCAGAATACCTCCAGAAAGAGACTAAAGGGAGC 4064
 901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917
 |||||
 4065 CACCCAGCGCTCAGTGGGAAGTGGACTTCACTGAGGTAAAGCCGGCTAA 4114
 917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934

```

|||||
4115 ATACGAAACAATATCTATTGGTTTGTAGACACCTTTTCAGGATGG 4164
934  aGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950
4165 TAGAGGCTTATCCTACATAAAAGAGACTTCAACCGTGGTGGCTAAGAAA 4214
951  IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967
4215 ATACTGGAGGAATTTTCCAGATTGGAATACCTAAGGTAAAGTATAGGGTC 4264
967  rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984
4265 AGACAATGGTCCAGCTTTCGTGTCGCCAGTAAGTCAGGAGCTGGCCAA 4314
984  leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000
4315 TATTGGGATTGATTGAAACTGTCATTGTGCATACAGACCCCAAGCTCA 4364
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017
4365 GGACAGGTAGAGAGATGAATAGAACCATTAAGAGAGACCTTACCAATT 4414
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034
4415 CACCACAGAGACTGGCATTAATGATTGGATGGCTCTCCTGCCCTTTGTGC 4464
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050
4465 TTTTAGGGTGAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATATAA 4514
1051 LeuLeuTyrGlyGlyProProLeuAlaGluIleAlaPheAlaHisSe 1067
4515 TTGCTCTACGGGGACCCCGCTTGGCAGAAATTCCTTTGCACATAG 4564
1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084
4565 TGCTGATGTGCTTTTCCAGCCTTGTCTCTAGGCTCAAGCGCTCG 4614
1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSerGly 1100
4615 AGTGGTGAGGACGCGGTGAAGACAGCTCCGGAGGCGCTACTCAGGA 4664
1101 GlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTyrVa 1117
4665 GGAGACTTGCAAGTTCCACATCGCTTCCAGTTGGAGATTTCAGTCTATG 4714
1117 lArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProTyrL 1134
4715 TAGACGCCACCGTCGAGGAACCTCGAGACTCGGTAGAAGGACCTTATC 4764
1134 euValLeuLeuThrThrProThrAlaValLysValGluGlyIleProThr 1150
4765 TCGTACTTTTGACACACCAACGCGTGTGAAGTCGAAGG..... 4804
1151 TrpIleHisAlaSerHisValLysProAlaProProAspSerGlyTr 1167
4804 ..... 4804
1167 pArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgLeuV 1184
4805 .....AATPCCCTTAAGCTTCGCCCTCATCGCGTGG 4835
1184 aProTyrSerAsnAsnSer 1191
4836 TTCCTTACTCTGTCAATACTCC 4858

```

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT.AA174883

seq_documentation_block:

ID AAT74883 standard; cDNA: 7393 BP.

XX

AC AAT74883;

```

XX 09-FEB-1998 (first entry)
DT Porcine retrovirus cDNA (defective).
XX
DE Retrovirus; porcine; GAG protein; POL protein; ENV protein;
XX xenotransplantation; Infectious; provirus; organ transplant; donor;
KW activated virus; PCR; ss.
KW Porcine retrovirus.
XX
XX Key Location/Qualifiers
FT CDS 598..2172
FT /*tag= a
FT mat_peptide 598..2169
FT /*tag= b
FT /*note= "putative GAG protein"
FT CDS 2320..4737
FT /*tag= c
FT /*note= "putative POL coding region (partial) as
FT described in the specification"
FT mat_peptide 2320..3522
FT /*tag= d
FT /*note= "putative POL protein (partial)"
FT mat_peptide 3516..4328
FT /*tag= e
FT /*note= "putative POL protein (partial)"
FT mat_peptide 4332..4748
FT /*tag= f
FT CDS 4738..6725
FT /*tag= g
FT /*note= "putative ENV coding region (partial) as
FT described in the specification"
FT mat_peptide 4752..6722
FT /*tag= h
FT /*note= "ENV protein (partial)"
XX
XX WO9721836-A1.
XX
XX 19-JUN-1997.
XX
XX 13-DEC-1996; 96WO-US19680.
XX
XX 14-DEC-1995; 95US-0572645.
XX (GEO ) GEN HOSPITAL CORP.
XX
XX Fishman JA;
XX
XX WPI: 1997-332804/30.
XX P-PSDB; AAW32091-W32095.
XX
XX New nucleic acid from porcine retroviruses - used for detecting
XX viruses in transplant or other tissue and for assessing risk of
XX transmitting infection to graft recipient
XX
XX Claim 16; Fig 2; 128pp; English.
XX
XX This cDNA sequence represents a defective purified swine retrovirus
XX found in PK-15 cells containing the putative coding regions for viral
XX GAG, POL and ENV proteins. There are a few in frame stop codons and
XX apparent frame shifts in the given coding sequence which alter features
XX of the translation. This sequence and PCR fragments generated from the
XX presence of porcine retroviruses prior to xenotransplantation.
XX Transplantation can increase the likelihood of retroviral activation
XX intact and infectious proviruses are present. The porcine retroviral
XX sequence can be used to generate probes to determine the level (e.g.
XX copy number) of intact (i.e. potentially replicating) porcine provirus
XX sequences in a strain of xenograft transplantation donors. It can be
XX used to detect mutations, genetic lesions or viral recombinants and to
XX determine the histological localisation of activated retrovirus. Using
XX Polymerase Chain Reaction DNA Quantitation (PQO) on blood mononuclear

```

CC cells, infectivity titration and susceptibility testing can be
 CC performed. Ultimately animal donors without intact porcine retroviral
 CC sequences or a lower copy number of viral elements could be selected.
 XX
 SQ Sequence 7393 BP; 2004 A; 1801 C; 1896 G; 1692 T; 0 other;

alignment_scores:

Quality: 4489.50 Length: 1191
 Ratio: 5.033 Gaps: 2
 Percent Similarity: 74.895 Percent Identity: 74.643

alignment_block:

US-09-171-553B-5 x AAT74883 ..

Align seq 1/1 to: AAT74883 from: 1 to: 7393

```

1  GLyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17
2173 GGGAGAGCGGGTTCCGACCCCTCCCGAGCCCGAGGGTAACCTTTGAAGGT 2222
17  lGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34
2223 GGAGGGGCACACAGTTGAGTTCTCGTTGATACCGGAGCGGACATTCAG 2272
34  alLeuLeuGlnProLeuGlyLysLeuLysGluLysLysSerTrpValMet 50
2273 TGCTGCTACAAACCATAGGAAACTAAAGAAAAAAATCCTCGGTGATG 2322
51  GLyAlaThrGlyGlnArgGlnTyrProTrpThrArgArgThrValAs 67
2323 GGTGCCACAGGCAAGCGAGTAGTATCCATGGACTACCCGAAGAACCGTTGA 2372
67  pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysP 84
2373 CTGGGAGTGGGAGCGGTAAACCCACTCGTTCTGGTCATCCCTGAGTGGC 2422
84  roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100
2423 CAGTACCCCTCTAGGTAGAGACTTACTGACCAAGATGGGAGCTCAAAAT 2472
101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIle 117
2473 TCTTTTGAACAGGAAGACAGCAAGTCTGTGAATAACAACCCATCAC 2522
117  rValLeuThrLeuGlnLeuAspGluTyrArgLeuTyrSerProGlnV 134
2523 TGTGTTGACCCCTCAATTAGATGATGAATATCGACTATATTTCTCCCAAG 2572
134  allysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150
2573 TAAAGCCTGATCAAGATATACAGTCTCGTTGGAGCAGTTTCCCAAGCC 2622
151  TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnVa 167
2623 TGGGCAGAAACCGCAGGGATGGGTTTGGCAAGCAAGTTCGCCACAGGT 2672
167  lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184
2673 TATTCAACTGAAGCCAGGTGCTACACAGTATCAGTCAGACAGTACCCCT 2722
184  euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200
2723 TGAGTAGAGAGGCTCGAGAAGGAATTTGGCCGCATGTTCAAGATTATC 2772
201  GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217
2773 CAACAGGGCATCTAGTCTCTGTCCTCAATCCCTTGAATACTCCCTGCT 2822
217  uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234
2823 ACCGGTTAGGAAGCCCTGGGACCAATGATTATCGACAGTACAGGACTTGA 2872

```

```

234  rGluValValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250
2873 GAGAGTCAATAAAGGGTGCAGGACATACACCAACGGTCCCGAACCTT 2922
251  TyrAsnLeuLeuSerAlaLeuProGluArgAsnTrpTyrThrValIle 267
2923 TATAACCTCTTGAGCGCCCTCCCGCTGAACGGAACCTGGTACACAGTATT 2972
267  uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284
2973 GGACTTAAAGATGCTTCTTCTGCTGAGATTACACCCCTAGCCCAAC 3022
284  roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
3023 CACTTTTACCTTCGAATGGAGATCCAGGTACGGGAAGAACCGGCGAG 3072
301  LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317
3073 CTCACCTGGACCCGACTGCCCAAGGGTTCAAGAACTCCCGACCATCTT 3122
317  eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
3123 TGACCAAGCCCTACACAGGACCTGCCCAACTTCAGGATCCACACCCCTC 3172
334  InValThrLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350
3173 AGGTGACCCCTCTCCAGTACGTGGATGACCTGCTTCTGGCGGAGCCACC 3222
351  LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAs 367
3223 AAACAGGACTGCTTAGAAGGTACGAAGCCACTACTGCTGGAATTGCTGA 3272
367  pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
3273 CTTAGCTACAGAGCTCTGCTAAGAGGCCAGATTTCGAGGAGAGAGG 3322
384  alThrTyrLeuGlyTyrSerLeuArgGlyGlnArgTrpLeuThrGlu 400
3323 TAACTACTTGGGGTACAGTTTGGCGGCGGCGAGCGATGGCTGACGAG 3372
401  AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysGl 417
3373 GCACGGAAGAAACTGTAGTCCAGATACCGGCCCAACACACAGCCAAACA 3422
417  nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
3423 AGTGAGAGAGATTTTGGGACAGCTGGATTTTGCAGACTGGGATCCCGG 3472
434  lypPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450
3473 GGTTCGAGACCTTAGCAGCCCACTTACCCTTAACCAAGAAAGAAAGGG 3522
451  GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467
3523 GGTTCG..... 3527
467  sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484
3527 ..... 3527
484  heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500
3527 ..... 3527
501  GlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLe 517
3527 ..... 3527
517  uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaLav 534
3527 ..... 3527
534  alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550

```

3527 3527
 551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567
 3527 3527
 567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuT 584
 3527 3527
 584 hrGluArgValThrPheAlaProProAlaLeuAsnProAlaThrLeu 600
 3527 3527
 601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617
 3527 3527
 617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634
 3527 3527
 634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650
 3527 3527
 651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667
 3527 3527
 667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684
 3527 3527
 684 euThrGlnAlaLeuArgLeuAlaGluLysSerIleAsnIleTyrThr 700
 3527 3527
 701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717
 3527 3527
 717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734
 3528CTTACCTCAGCAGGAGGAGAAATAAGAACAAAGAGG 3564
 734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750
 3565 AAATTCAGACCTATTAGAACCTTACATTGGCCAAAGAGCTAGCTATT 3614
 751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767
 3615 ATACACTGTCTGGACATCAGAACGCCAAGACTCATATCTAGAGGGAA 3664
 767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784
 3665 CCACATGGCTGACCGGGTTGCCAAGCAGGCAGCCAGCGCTGTTAAACCTC 3714
 784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800
 3715 TGCCATAAATAGAAACCCCAAGGCCCAAGAACCCAGACAGACAGTACACC 3764
 801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817
 3765 CTAGAACACTGGCAGAGATATAAAGATAGACACAGTCTCTCAGACTCC 3814
 817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834
 3815 GGAGGGACCTGCTATACCTCATATGGGAAGGAATCCTGCCCCACAAAG 3864
 834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850

3865 AAGGGTTAGATATGTCCAACAGATACATCGTCTAAACCACCTAGAACT 3914
 851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867
 3915 AAACACCTGCAGCAGTTGTCAGAACATCCCTTATCATGTCTGAGGCT 3964
 867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884
 3965 ACCAGAGTGGCTGCTCGTGGTCAAAACATTGTGTGCCCTGCCAGCTGG 4014
 884 alAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900
 4015 TTAATGCTTAATCCTTCCAGATACCTCCAGAAAGAGACTAAGGGAAAGC 4064
 901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917
 4065 CACCCAGGGGCTCACTGGGAAGTGACTTCACTGAGGTAAGCCGGCTAA 4114
 917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934
 4115 ATACGGAAACAAATATCTATTGGTTTTTGTAGACACCTTTTCAGGATGG 4164
 934 alGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950
 4165 TAGAGCTTATCCTACTAAAAAGAGACTTCAACCGTGGTGGCTAAGAA 4214
 951 IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967
 4215 ATACTGGAGGAAATTTTCCAAGATTGGAAATACCTAAGGTAATAGGCTC 4264
 967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984
 4265 AGACAATGGTCCAGCTTTGCTTGCACAGTAAGTCAGGACTGGCCAAAG 4314
 984 leuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000
 4315 TATTGGGATTGATTGAAACCTGCATTGTGCATACAGACCCCAAGCTCA 4364
 1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017
 4365 GCACAGGTAGAGAGGATGAATAGAACCATTAAGAGACCTTACCATAAT 4414
 1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034
 4415 GACCACAGACTGGCATTAATGATTGGATGGCTCTCTCGCCCTTTGTGC 4464
 1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050
 4465 TTTTAGGGTGAGAACACCCCTGGACAGTTTGGGCTGACCCCTATTAAA 4514
 1051 LeuLeuTyrGlyGlyProProProLeuAlaGluIleAlaPheAlaHisSe 1067
 4515 TTGCTCTAGGGGGACCCCCCGCTGGCAGAAATTCCTTTGCACATAG 4564
 1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084
 4565 TGCTGATGTGCTGCTTCCAGCCCTTGTCTCTAGGCTCAAGGCGCTCG 4614
 1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSerGly 1100
 4615 AGTGGGTAGGCAGGAGCGTGGAGCAGCTCCGGGAGCCCTACTCAGGA 4664
 1101 GlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTyrVa 1117
 4665 GGAGACTTGCAAGTTCCACATCGCTTCCAAAGTTGCAGATTTCAGTCTATGT 4714
 1117 lArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProTyrL 1134
 4715 TAGAGCCACCCGTCGAGGAAACCTCGAGACTCGTAGAAGGACCTTATC 4764
 1134 euValLeuLeuThrThrProThrAlaValLysValGluGlyIleProThr 1150
 4765 TCGTACTTTTGGACCACACACGCGCTGTGAAGTCGAAGG..... 4804

151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProGlnVa 167
167 lileGlnLeuLysAlaSerAlaThrProValSerValArgGlnThrProL 184
1788 GGTAGAAGCTTAAGCTGATGATCACCCTCCATCTCAGTAAGCAATACCCCA 2837
184 euSerArgGluAlaArgGluGlyIleThrProHisValGlnArgLeuIle 200
2838 TGACGAAGAGCTAAGAGGGGATCCGACCTCATATTCGGAGGCTGCTA 2887
201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217
2888 GACCAAGGAATTTAGTGGCTGTGAGTCCCTGGAACATCCACATTTT 2937
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234
2938 GCCAGTAGCAAGGACGAGGACCAATGATTATCGCCCGGTCAAGACCTCC 2987
234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250
2988 GGAAGATTAAAGGGTCTCGACATTCACCTACAGTCCCGAACCCG 3037
251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValLe 267
3038 TACAACTTACTAAGCTCTCTCCACCCGAGAGAACCTGGTATACGGTCT 3087
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284
3088 GGATTAAAGATGCTTCTTTTGGCTGCGTCTGCACCCCAAGAGTCAAC 3137
284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
3138 TTCGTCTTGGCTTGAATGGAGAGACCCAGAGGGCGGACAGACTGGTCAA 3187
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317
3188 TTAACCTGGACTAGACTACCAAGGGGTTCAAAACTCCCTACCTCCCTGT 3237
317 easpGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
3238 TGACGAAGCCCTCCATCGGGATTAGCACCTTTCGGCTCAAAACCCCC 3287
334 InValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350
3288 AGCTTACCCTACTGCAGTACGTAGATGATCTTTAATCGCAGCGCCTCA 3337
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuSerAs 367
3338 AAAGAGCTATGTCAACAGGGGACTGAGAGACTCCTCCACAGAACTGGGGA 3387
367 pleuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
3388 YTTGGGGTATCAGTTTCGGCTAAAGGCACAAATTGTCTCAACTCAGG 3437
384 alThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400
3438 TAATCTACTGGGGTATCTTTGCGAGGAGGTAAAGATGCTCCTCAGTCAA 3487
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG 417
3488 GCCCGGAAAAAGACTGTCTATGATGATTCACCCGCCCAACACACCCAGGCA 3537
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
3538 GTACGCTGACTTCTGGGGACTGCTGGCTTTGTAGACTCTGGATTCCAG 3587
434 lypheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450
3588 GCTTTGCGACCTAGCAGCACCCTTATATCCCTTACTAGGGAAGGAATT 3637
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467

3638 CCTTTGAATGGAAGAAACACCAAGAGCTTTTGAGGCTATCAAAATC 3687
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484
3688 GTCTCTAATGACTGCCCTCGCGCTAGCATTACCAGACTTGACTAAGCTCT 3737
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500
3738 TCGTCTCTATATGTGGACGAGAGCGGCGCATAGCCAGAGGGGTGCTGACA 3787
501 GlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLe 517
3788 CAAGCACTAGGACCTTGGAGAGACCGGTAGCTATCTGTCAAAAAAAT 3837
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaAlaV 534
3838 GGATCCCGTGTAGTGGATGCCACATCTCTGAAAGCTATTGCAGCAG 3887
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550
3888 TAGCCCTGTGATCAAAAGATGCTGATAAACTGACAATGGAGCAGAGGTG 3937
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567
3938 ACTGTCTGGCCCTCATGCCCTGGAGAGTATTGTGGCGCAGCCACTGA 3987
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeu 584
3988 TAGATGGATGACTAATGCCCGGATGACACACTATCAGAGTCTGCTGCTAA 4037
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600
4038 ATGACCGGGTAACTTTGCCCCCTTCTCAACCCAGCCACCCCTC 4087
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617
4088 CTCCCTCTAACGAATGATTCCTGCCAGTACATCATGTCGACAGCATCTCT 4137
617 uileGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634
4138 GGCTGAAGAAATGGGACAGAAAGACCTGACTGACCAACCTGGCCCTG 4187
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650
4188 GA...GCTCTAGTGTACACGATGGCAGCAGTCTTCTGATTGAGGA 4234
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667
4235 AAGCGAGGGCTGGAGCTGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4284
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684
4285 AAGTGCCTTGCCTGAAGAACTTCGCGACAGAAAGGCTGAACATCATAGCG 4334
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700
4335 TTACACAGGCCCTCCGAGAGGCCGAGGTAAGATCATTAATATTACACT 4384
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717
4385 GACAGCGCTATGCTTTGCTACTGCACATATCCATGGGGCCATCTACAG 4434
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734
4435 GCAGCAGGGCTTATGACTTACGCGGTAAAGACATCAAAACAAAGAAAG 4484
734 lulleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750
4485 AAATTTGCGCCCTTAGAAGCCATCATGCGCCCAAGAGGTAGGCATC 4534
751 ileHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgLys 767

4535 ATCCACTGTCCTCCGCTCACCAGAAAGGAGGAGGACTTAGTGGCCCAAGGGCAA 4584
767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784
4585 CCGGATGGCAGACTCAGTACAAACAGGTGTCTCAGGGGCCATGATCT 4634
784 euProIleIleGluThrPro...LysAlaProGluProArgArgGlnTyr 799
4635 TAACGTAAAGAAATCCGTCCAAAAGCCCTGAG..... 4669
800 ThrLeuGluAspPrpGlnGlnIleLysLysIleAspGlnPheSerGluTh 816
4670GATGAAACTATGATATAAAAGAACTATTTGGAC 4704
816 rProGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisL 833
4705 TAGTGATCCCTCCCATCTTTTCGAAGGAAAAATAGACTTCACACCCG 4754
833 ysGluGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGly 849
4755 AGGAAGGATAAAATTTGTCAAAGGACTACACCACTTACCCACCTGGGA 4804
850 ThrLysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuAr 866
4805 GTTGAAAAATGATGAGACTGATTAAGAGTCCCGGTATCAAGTCCTAA 4854
866 gLeuProGlyValAlaAspSerValValLysHisCysValProCysGlnL 883
4855 CTGGAAGTCTAGTGGCTCAAAAGATTAATAACTCTCGAAAGCGTGTGCAT 4904
883 euValAsnAlaAspProSerArgIleProProGlyLysArgLeuArgGly 899
4905 TCACATATGCACTAAACCTACAAAGACCTGGAAGAGACACACGGGA 4954
900 SerHisProGlyAlaHisTrpGluValAspPheThrGluValLysProAl 916
4955 GACCGTCTCGAGTGTATCGGAGGTGGATTTTACTGAAGTTAAACCCGG 5004
916 aLysTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyT 933
5005 AATGATGGTAATAGTATCTGTAGTATTTGTAGACACCTTTTCAGGAT 5054
933 rpValGluAlaTyrProThrLysLysGluThrSerThrValValAlaLys 949
5055 GGGTTGAAGCGTTCCTCCACTAAACCTAGACTGCCAGATTTGTGGCAAG 5104
950 LysIleLeuGluGluIlePheProArgPheGlyIleProLysValIleGl 966
5105 AAGATTTTGAAGAAATCCCTGCCAAGATATGGAGTACCTAAGGTAATCG 5154
966 ySerAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaL 983
5155 GTCCGACATGACACCGCTTTGTGCCAGGTAAAGTCAGGATTTGGCCA 5204
983 ysIleLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSer 999
5205 CTCAGTTGGCGATTGATTGGAATTTACATTGTCTTACCCCTCAAGC 5254
1000 SerGlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLy 1016
5255 TCAGGACAGGTAGAGAGGATGAATAGAACATTAAAGGACACCTTGACTAA 5304
1016 sLeuThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheV 1033
5305 ATTTGGCCATCGAGACCGGGGAAAGACTGGGTGGCTCTCTCCCTCCCTTG 5354
1033 allLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyr 1049
5355 CGCTCTCCGCGCCGGAACACCCCGGAGGTTTGGGCTCAGCTCCTTTT 5404
1050 GluLeuLeuTyrGlyGlyProProProLeuAlaGlu.....IleAl 1063
5405 GAAAGTCTGTGGGGACCTCCCTCTTTTAAATAAAGATGGTGGAAACATT 5454

1063 aPheAlaHisSerAlaAspValLeuLeuSerGlnProLeuPheSerArgL 1080
5455 GGTTCCTCCGATTCAGGCTCTGTCTTACCCTCTCT...TTGCTTATCAT 5501
1080 euLysAlaLeuGluTrpValArgGlnArgAlaTTrpLysGlnLeuArgL 1096
5502 TAAAGCCCTGAAAGTAATTAGGACCCCAAAATTTGGGACACGCTAAAGACG 5551
1097 AlaTyrSerGlyGlyAspLeuGlnValProHisArgPheGlnValGlyAs 1113
5552 GCCTACACCCCGAGGACACCCAGCATCCGCCACGAGTTCAGGTCGGGA 5601
1113 pSerValTyrValArgHisArgAlaGlyAsnLeuGluThrArgTrpL 1130
5602 CCAAGTCTTGGTCAGACGACATCGAACCGGTAGCTTGAACACCGTGA 5651
1130 ysGlyProTyrLeuValLeuLeuThrThrProThrAlaValLysValGlu 1146
5652 AGGGACCTTATTTAGTGTACTAACAACTCTTACGGCAGTGAAGTTGAC 5701
1147 GlyIleProThrTrpIleHisAlaSerHisValLysProAlaProProPr 1163
5702 GGGATTGCTCTCTGGATCCACGCTTCACACGTCAGAGGGCCCGCAGTCA 5751
1163 oAsp.....SerGlyTrpArgAlaGluLysThrGluAsn 1175
5752 AGATGAAGAAACCCAGGAGACAACTGGGCTGTGGAGCCACTGATAACC 5801
1175 rLeuLysLeuArgLeuHisArgLeuValProTyrSerAsnAsnSer 1191
5802 CTCTTAAGCTTCGCTGCGTCGAGGAGGCCCTC.....CATCAC 5842
1192 ProGly 1193
5843 CCTGGA 5848

FT	protein_bind	7885..7914
FT	/tag= p	
FT	/note= "Binding site for transcription factors AP-1/GATA, c-Myd and Ap-1"	
FT	protein_bind	7928..7953
FT	/tag= q	
FT	/note= "Binding site for transcription factors E47 and ETS-1"	
FT	CAT_signal	7964..7968
FT	/tag= r	
FT	protein_bind	7985..7997
FT	/tag= s	
FT	/note= "Binding site for transcription factor GATA"	
FT	CAT_signal	8003..8007
FT	/tag= t	
FT	protein_bind	8024..8036
FT	/tag= u	
FT	/note= "Binding site for transcription factor GATA"	
FT	CAT_signal	8042..8046
FT	/tag= v	
FT	protein_bind	8049..8060
FT	/tag= w	
FT	/note= "Binding site for transcription factor AP-1/CREB"	
FT	TATA_signal	8075..8090
FT	/tag= x	
FT	misc_feature	1..61
FT	/tag= y	
FT	/note= "R LTR domain"	
FT	misc_feature	8127..8135
FT	/tag= z	
FT	/note= "Polya downstream element"	
FT	polya_signal	8161..8166
FT	/tag= aa	
PN	W09740167-A1.	
XX	30-OCT-1997.	
PD	18-APR-1997;	97WO-GB01087.
XX	10-FEB-1997;	97GB-0002668.
PR	19-APR-1996;	96GB-0008164.
XX	(IMUT-) IMUTRAN LTD.	
PA	(QONE-) Q-ONE BIOTECH LTD.	
XX	Gabraith DN, Haworth C, Lees GM, Smith KT;	
P1	WPI: 1997-535851/49.	
DR	P-PSDB: AAM39271; AAM39272; AAM39273.	
XX	Polyonucleotide encoding porcine retrovirus expression product -	
PT	xeno-transplantation	
PS	Claim 4; Fig 3; 69pp; English.	
CC	This DNA sequence encodes the porcine retrovirus (PoEV) virion core polypeptide (GAG), polymerase (POL) and envelope (ENV) proteins and also includes the Long Terminal Repeat (LTR). These proteins can be used to develop viral vaccines, antisense nucleic acids, ribozymes and other antiviral agents. They can also be used in xeno-transplantation technology and as diagnostic tools.	
SQ	Sequence 8209 BP; 2168 A; 2064 C; 2154 G; 1823 T; 0 other;	

```
alignment_scores:      Length:      656
                      quality: 3515.00
                      Ratio:   5.358
                      Gaps:     0
Percent Similarity: 100.000  Percent Identity: 100.000
```

```

alignment_block:
US-09-171-553B-6 x AAV09700      ..
Align seg 1/1   to: AAV09700   From: 1   to: 8209

```

1	MeHhProThrlenseuSrArgrHlsleuProHrrArvglyGlyVgLuPr	17
5620	ATGCAATCCACGCTTAAAGCCGGCCCACTCCCGCACTCGGGGTGGAGACC	5659
17	oLysArGLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuT	34
5670	GAAAGACGTAGAAATCCCTTAAGCTTGCGCTCCATCCCTGGATCTCTTA	5719
34	hrlneuHrlleuHrrProGlnAlaSerSerLysArgLeuIleAspSerSer	50
5720	CTGTAAACAATAACTCCCGCAGGCACAGTAAGAACGCTTATGACAGCTCG	5769
51	AspProHisArgProLeuSerLeuHrrTrpLeuIleIleAspProAspTh	67
5770	AACCCCAATAGACCTTATACCTTACCTACGTGCGGATATTAGAACCTGATAC	5819
67	rGlyValThrValAsnSerThrArgIlyValAlaProArgIlyHrrTrpT	84
5820	GGGTGTCACTGTAAATAGCACTCGAGGTGTGCTCCTTAGAGCACTGCTG	5869
84	rPrProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer	100
5870	GGGCTGACGACGCTTCTCTCTCCGATTGATTAAACCCGCTGTAAAAAC	5919
101	ThrProProAsnLeuValArgSerTrgIlyPheTrgCysCysProGlyTh	117
5920	ACACTCCCAACTACTGCTGATGATTATGGTGTCTATGCTGCCAGGCAC	5969
117	rGlyLysGlyLysTrgCysGlyGlySerGlyLysSerPheCysArgArgT	134
5970	AGGAAAGAGAAATACTGTGGGGGTGTGGGAAATCTTGTGTAGAAAT	6019
134	rPserCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu	150
6020	GGAGCTGCGTCACTCCAAACGATGGAGACTGGAATGGCCGATCTCTCTC	6069
151	GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTrgLy	167
6070	CAGGACCGGGTAAATAATCTCTTTGCAATTCGCCGCCGCAAGTACA	6119
167	sMetMetLysLeuTrgLysAspLysSerCysSerProSerAspLeuAspT	184
6120	AATGATGAACAATATAAAGATAGAGCTGCTCCCATAGACTTACTGATT	6169
184	yrLeuLysIleSerPheHrrGluArgLysThrGlyLysTrgSerLysVal	200
6170	ATCTAAAGATAAGTTTCACTGAAAGGAAACAGAAATATTCAAAAGTG	6219
201	AspLysTrpTrgIleuGlyAsnSerPheLeuLeuTrgIlyGlyGlyAla	217
6220	GATAAATGTGATACGCTGGGGAAATAGTTTATTATATGGCGGGGAGC	6269
217	agLysSerThrLeuThrIleArgLeuArgIleGluHrrGlyHrrGluProP	234
6270	AGGGTCACTTAAACATTTCGCTTAGATAGAGCGGGGACAGAACCC	6319
234	roValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProProAla	250
6320	CTGTGGGAATGGGACCGATAAAGTACTGGTGAAACAGGGGGCCCCGGCC	6369
251	LeuGluProProHisAsnLeuProValProGlnLeuHrrSerLeuArgPr	267
6370	CTGGAGCACCGCATTAACCTTGCGGGGCCCAATTAACCTGCTCGCCGCC	6419
267	oAspIleHrrGlnProProSerAsnSerHrrHrrGlyLeuIleProHrrA	284
6420	TGACATACACACCGCGCTTAGACACAGTACCACTGAGATTGATCTCTACA	6465

```

284 snThrProArgAsnSerProGluValProValIysThrGlyGlnArgLeu 300
|||||
6470 ACACGGCTAGAAACTCCCGAGTGTCTCTTACAGACAGAGACAGACACTC 6519
301 PheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspProAs 317
|||||
6520 TTCAGTCTCATCCAGGAGCTTCCAAAGCATCACTCACAGACGACCTTGA 6569
317 PAlaThrSerSerCysTrpLeuGlySerLeuSerSerGlyProProTyrTyrG 334
|||||
6570 TGGCCTCTCTCTTGTGGCTGTCTATCCCTCAGGCGCTCTTATTATG 6619
334 IuGlyMetAlaLysGluArgLysPheAsnValThrLysGlnHisArgAsn 350
|||||
6620 AGGGGATGGCTAAAGAAAGAAATTCATATGTGACCAAGAGCATTAAGAA 6669
351 GlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSerG 367
|||||
6670 CAATGTACATGGGGGTCCGAAATTAAGCTTACCCCTCAAGTCTTCCG 6719
367 yLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeuCys 384
|||||
6720 GAAGGGACATGCATAGAGAAACCTCCCTCCACCAACACCTTTGCT 6769
384 ySerThrValValItyrGluGlnAlaSerGlnAsnGlnItyrLeuValPro 400
|||||
6770 ATAGTACTGTGTTTATGACACAGGCTCAGAAATCAGTATTAGTACCT 6819
401 GLyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValSe 417
|||||
6820 GGTATTAACAGTGTGGGTGCATATCTGCTTAACCCCTGTGTTTC 6869
417 rThSerValPheAsnGlnSerLysAspPheCysPheValMetValGlnIle 434
|||||
6870 CACCTCAGTCTCAACCAATCAAGATTTCTGTCTATGTCATGTCACAAATG 6919
434 aLProArgValItyrTyrHisProGluGlnValValLeuAspGluTyrAsp 450
|||||
6920 TCCCGCAGTGTACTACCATCTGAGSAACTGCTCTGTATGAATATGAC 6969
451 TyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAla 467
|||||
6970 TATCGGTATTAACGACCAAAAGAAACCCGTATCCCTTACCTAGCTGT 7019
467 lMetLeuGlyLeuGlyThrAlaValGlyValIleGlyThrAlaAla 484
|||||
7020 AATGCTCGATTAGGAGGCGCTTGGCTTAGGAACAGGAGACGCTGCC 7069
484 euIleThrGlyProGlnGlnLeuGlnLysGlyLeuGlyLeuHisAla 500
|||||
7070 TGATCAGACAGACACAGCAGCTAGAGAAAGACTTGGTGAAGTAAATGCG 7119
501 AlametThrGluAspLeuArgAlaLeuLysGlySerValSerAsnLeu 517
|||||
7120 GCCATGACAGAAATCTCCAGGCTTAAAGAGCTGTTTGCAACCTAGA 7169
517 uGluSerLeuThrSerLeuSerGlnValValLeuGlnAsnArgArgL 534
|||||
7170 AGAGTCCCTGACTCTTGTCTGAGTGTCTTACAGAACCGAGGAGAT 7219
534 euAspLeuLeuPheLeuArgGluGlyLysLeuGlyLysAlaAlaLeuLys 550
|||||
7220 TAGATCTGCTGTCTTAAAGAGAGGTGGTATGTGACGCTTAAAGAA 7269
551 GluGlySerCysPheTyrValAspHisSerGlyAlaIleArgAspSerMet 567
|||||
7270 GAAATGCTCTTCTATGTAGATCACTCAGAGCCATCAGAACCTCCATGAA 7319
567 nLysLeuArgLysLysLeuGlnArgArgArgArgGluArgGluAlaAsp 584
|||||
7320 CAAGCTTAGAAAGAGTTAGAGAGCGCTCGAAGGAAAGAGAGCTGACC 7369
584 lngIyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThrThrLeu 600

```

```

|||||
7370 AGGGGTGTTGAAAGATGTTCACAGGCTCTCTTGCATGACACACCTG 7419
601 LeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuThrVa 617
|||||
7420 CTTTGTGCTGTGACGGGCGCTTGTAGTCTGCTGTCTTACTTACAGT 7469
617 lGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgVal 634
|||||
7470 TGGCCTTCTCTTAAATTAATAGTGTGTGCTTCTTGTAGAGACAGTGA 7519
634 eAlaValGlnIleMetValLeuArgGlnGlnItyrGlnGlyLeuLeuSer 650
|||||
7520 GTCCAGTCCAGATCATGTACTTAGGCAACAGTACCAAGGCTTCTGAGC 7569
651 GlnGlyGluThrAspLeu 656
|||||
7570 CAAGGAGAAACATGACCTC 7587

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:AAV09699
seq_documentation_block:
ID AAV09699 standard; DNA: 8196 BP.
XX
AC AAV09699;
XX
DT 19-MAY-1998 (first entry)
XX
DE Porcine retrovirus DNA encoding, GAG, POL and ENV.
XX
KW Porcine retrovirus; POEV; POL protein; ENV protein; GAG protein;
KW vaccine; diagnosis; xenotransplantation; prophylactic; therapeutic; ds.
XX
OS Porcine retrovirus.
XX
FH Key
FH Location/Qualifiers
FT 576..2126
FT /*tag= a
FT /product= GAG protein
FT /note= "viral core polypeptide"
FT 2143..5733
FT /*tag= b
FT /product= POL protein
FT /note= "polymerase peptide sequence as given in
FT specification"
FT 5606..7576
FT /*tag= c
FT /product= ENV protein
FT /note= "envelope protein"
XX
PN MO9740167-A1.
XX
PD 30-OCT-1997.
XX
PE 18-APR-1997; 97WO-GB01087.
XX
PR 10-FEB-1997; 97GB-0002668.
PR 19-APR-1996; 96GB-0008164.
XX
PA (IMOT-) IMOTRAN LTD.
PA (OONE-) O-ONE BIOTECH LTD.
XX
PI Galbraith DN, Haworth C, Lees CM, Smith KT;
XX
DR WPI: 1997-535851/49.
XX
PT Polynucleotide encoding porcine retrovirus expression product -
PT useful to develop products for use in vaccines, diagnosis and
PT xeno-transplantation
XX
PS Claim 4; Fig 2; 6pp; English.
XX
CC This DNA sequence encodes the porcine retrovirus (POEV) virion core

```


CC polypeptide (GAG), polymerase (POL) and envelope (ENV) proteins. These
CC proteins can be used to develop viral vaccines, antisense nucleic acids,
CC ribozymes and other antiviral agents. They can also be used in
CC xeno-transplantation technology and as diagnostic tools.
XX

SQ Sequence 8196 BP; 2165 A; 2061 C; 2147 G; 1820 T; 3 other;

```
alignment_scores:
    Quality: 3509.00
    Ratio: 5.357
    Percent Similarity: 99.848
    Length: 656
    Gaps: 0
    Percent Identity: 99.848
```

```
alignment_block;
```

Align seg 1/1 to: AAV09699 from: 1 to: 8196

1	MEHISProThrlleuSerArGdVghISleuProThrArgIglYgluPr	17
5606	ATGCATCCCAACGTTAAAGCTGGCCACACTCCCGCTCGGGGTGGAGACC	5655
17	OLysArGlEuArGlIeProleuSerPheAlaSerIleAlaItrPheLut	34
5656	GAAAAGACTGAGATCCCTTAAGCTTCGGCTCCATCGCTGGTTCCTTA	5705
34	hrlEuThrlIethrProGlnAlaSerSerLysArGlEuIleasPserSer	50
5706	CTCTACATATACCTCCCGAGCGACAGTAATAACGCCCTTATAGACAGCTCG	5755
51	AsnProHISArGProleuSerLeuThrTlPleuIleIleasProAspTh	67
5756	AACCCCATAGACTTATTCCTTACCTGGCTGATATATGACCTGATAC	5805
67	rgLYaIhThrValasnsSerThrArgglYAlaIaProArGglYThrTrp	84
5806	GGGTGTCACTGTAAATAGACACTCGAGGTGTGCTCTAGAGGACCGT	5855
84	rPProGlnleuHISphecYSleuArGlEuIleasnProAlaValysSer	100
5856	GGCCTGAACCTGCATTTGTGCGCCGATGTATTAACCCGCTTAARACG	5905
101	ThrProProAsnleuValArGserThrGlyPheThrCYsCysProGlyTh	117
5906	ACACCTCCCAACCTAGTGCCTGATATGGCTCTATGTCTGCCACGACAC	5955
117	rgLlLysleuLysThrCYsGlyGlySerGlyLusSerPhecYSArGArG	134
5956	AGAAAGAGAGAAATACGTGGGGGTTCTGGGGAATCCTTCTAGAGAGAT	6005
134	IPserCYsValThrSerasnAspGlyAspTrPlyTrProlleSerleu	150
6006	GGACTCGCTACCTCCACACATGGAGACTGGAAATGGCGCATCTCTC	6055
151	GlnAspArGValLysPheSerPheValasnsSerGlyProGlyLysThrY	167
6056	CAGACCCGGGTAAATCTCCTTGTCAATCCGGCCCGGGAAGTACAA	6105
167	SMetLysleuLysTrLysAspLysSerCYsSerProSerAspLeuAspT	184
6106	AATGATATAACTATATATAAGATAAGAGCTGCTCCCAACAGACTTAGATT	6155
184	YrLeuLysIleSerPheThrGluArGlysthrgLYLysSerLysVal	200
6156	ATCTTAAGATAGATTTCATCGAAGGAAACAGGAAATATTCAAAGTG	6205
201	AspLysTrPlyrGluLeuGlyAsnSerPheLeuLeuThrGlyGlyGlyAl	217
6206	GATAAATGGTATAGCTGGGGAATAGTTTTTATTATATGCGGGGGGAGC	6255
217	aglSerThrLeuThrIleArGleuArGlieGluThrGlyhrgLuprOP	234
6256	AGGGTCACATTTAACATTCGCTTACGATATGAGACGGGGACAAGAACCC	6305

234	rowAlaMetGlyProAspIysValIleuAlaGluGlnGlyProProAla	250
6306	CTGTGGCAATGGGACCGCATAAATGACTGGCTGAACAGGGGGCCCCGGCC	6353
251	LeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuArgPr	267
6356	CTGGAGCCACCGCATTACTTCCGGTCCGCCATTAACTCCGTGGGGCC	6405
267	AspIleThrGlnProProSerAsnSerThrGlyLeuIleProThra	284
6406	TCACATTAACAGACGGCGCTAGCAACACTACCTGATGTAACTCTACCA	6455
284	snThrProArgAsnSerProGlyValProValIysThrGlyGlnArgLeu	300
6456	ACAGCCCTAGAAACTCCCGCAGGTTCCTCTTTAAGCAGGACAGAGACTC	6505
301	PheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspProAs	317
6506	TTCAAGTCTCAATCCAGGGAGCTTCCAAAGCCATCAACTCCACCGACCTGA	6555
317	PalThrSerSerCysThrLeuGlySerGlyLeuSerGlyProProTyrG	334
6556	TGCCACTTCTTCTGTGTGGCTTGTCTTATCCCTACGGCGCTCTCTTTATGC	6605
334	IuGIyMetAlaIysGluIuArgIysPheAsnValThrLysGluHisArgAsn	350
6606	AGGGGATGGCTAAAGAAAGAAATTCATATGTCACCAAGAGCATGTGAAT	6655
351	GlnCysThrPrpGlySerArgAsnIysLeuThrIleThrGluValSerG	367
6656	CATGTACATAGGGGGGTCCGAAATTAAGTTACCTCCCTACGTGAAGTTTCCGG	6705
367	YLYSGIYThnCysIleGlyLysAlaProProSerHisGlnHisLeuGlyST	384
6706	GAAAGGACATGCGATAGGAAAGCTCCGCCATCCACCAACACTTTGCT	6755
384	YrSerThrValValYrGluGlnAlaSerGluAsnGlnYrLeuValPro	400
6756	ATATGACTGTGGTTTATGTAGAGGGCTCAGAAATACAGTATTATGTACTCT	6805
401	GlyYrAsnArgPrPrPrPalaCysAnthrGlyLeuThrProCysValSe	417
6806	GGTTATTAACAGGCGTGGGCATGCAATCTGGGTTTAAACCCCTGTGTTC	6855
417	rThrSerValPheAsnGlnSerLysAspPheCysValMetValGlnIleV	434
6856	CACCTCAGCTTCAACCAATCCAAAGATTTCTGHTCATGGTCCAAATCG	6905
434	alProArgValYrYrHisProGluGluValValIleuAspGluYrAsp	450
6906	TCCCCCGAGTGTACTACATCTCTGGAGAAAGTGCTGTGATGAATATGAC	6955
451	YrArgYrAsnArgProLysArgGluProValSerLeuThrLeuAlaVal	467
6956	TATCGGTATAACCGACCAAAAGAAACCCTATCCCTTACCCTACTGTG	7005
467	IleThrLeuGlyLeuGlyYrThrAlaValGlyValGlyYrThrAlaAlaL	484
7006	AATGCTCGGATTATAGGAGCGCGCTGGTGGTATAGAAACAGGACAGCTGCC	7055
484	eulIeThrGlyProGlnGlnLeuGluLysGlyLeuGlyGlnLeuHisAla	500
7056	TCATCACAAGGACCAAGAGAGGTATAGAAAGAGACTGGTGACTACATTGGC	7105
501	AlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnLeuG	517
7106	GCATAGACAGAGAACTCTCCGAGCCTTAAAGAGACTGTGTAGCAACTACA	7155
517	uGluSerLeuThrSerLeuSerGluValValIleuGlnAsnArgArgGlyL	534
7156	AAAGTCCCTGACTCTTTGTCTGAAGTGGTGTCTACAGAAACCGGAGGGGAT	7205

534 euAspLeuPheLeuAArgGluGlyLeuCysAlaLeuLysGlu 550
 |||||||
 7206 TAGATCTGCTGTTCTTAGAGAGAGGTGGTATGTGACGCTTAAGAA 7255
 551 GluCysCysPheTyrValAspHisSerGlyAlaLeuArgSerMetAs 567
 |||||||
 7256 GAATGTGCTGCTTATGATGATCACCAGAGGCAACAGACTCCATGAA 7305
 567 PLYSLeuAArgLysLysLeuGluAArgAArgAArgGluAArgLysAsp 584
 |||||||
 7306 CAAGCTTGAGAAAAAGTTAGAGAGCCGTGCAAGGAGAAAGAGGCTGACC 7355
 584 lnglYTrPheGluGlyTrPheAsnArgSerProTrpMetThrLeu 600
 |||||||
 7356 AGGGGTGGTTGAGAGATGGTTCAACAGGTGCTCCCTTGATGACCACTG 7405
 601 LeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuThyVa 617
 |||||||
 7406 CTTTCTGCTGACGGGCCCCCTAGTCTGCTGCTGCTGCTGCTGCTG 7455
 617 lglYProCysLeuLeuAsnArgPheValAlaPheValArgGluArgValS 634
 |||||||
 7456 TGGGCTTGCTGCTTAATTAATAGTTGCTGCTGCTGCTGCTGCTGCTG 7505
 634 eAlaValGlnIleMetValLeuAArgGlnIleTyrGlnGlyLeuLeuSer 650
 |||||||
 7506 GTGCAGTCCAGATCATGTACTTAGCAACAGTACCAAGGCTCTGAGC 7555
 651 GlnGlyGluThrAspLeu 656
 |||||||
 7556 CAAGAGAAACTGACCTC 7573

seq_name: /SIDS2/gcgdata/geneseq/geneseqr/NA1997.DAT.AAV09703

seq_documentation_block:

ID AAV09703 standard; DNA: 1974 BP.

XX AAV09703;

XX 20-MAY-1998 (first entry)

XX Porcine retrovirus Raji clone ENV DNA.

XX Porcine retrovirus; POEV; ENV protein; envelope protein; vaccine;

KW diagnosis; xenotransplantation; prophylactic; therapeutic; ds.

XX Porcine retrovirus.

XX Key Location/Qualifiers

FT CDS 1..1974

FT /*tag= a

FT /product= ENV protein

FT /note= "envelope protein"

XX MO9740167-A1.

XX 30-OCT-1997.

XX 18-APR-1997; 97MO-GB01087.

XX 10-FEB-1997; 97GB-0002668.

XX 19-APR-1996; 96GB-0008164.

XX (IMUT-) IMUPRAN LTD.

XX (QONE-) Q-ONE BIOTECH LTD.

XX Galbraith DN, Haworth C, Lees GM, Smith KT;

XX WPI: 1997-535851/49.

XX P-PSDB: AAW39274.

XX Polynucleotide encoding porcine retrovirus expression product -
 PT useful to develop products for use in vaccines, diagnosis and

PT xeno-transplantation
 XX
 PS Claim 6; Fig 4; 69pp; English.
 XX
 CC This sequence encodes the porcine retrovirus envelope (ENV) protein
 CC isolated from the human cell line Raji. Such viral proteins can be used
 CC to develop viral vaccines, antisense nucleic acids, ribozymes and other
 CC antiviral agents. They can also be used in xeno-transplantation
 CC technology and as diagnostic tools.
 XX
 SQ Sequence 1974 BP; 534 A; 489 C; 483 G; 468 T; 0 other;

alignment_scores:
 Quality: 3455.00 Length: 658
 Ratio: 5.291 Gaps: 2
 Percent Similarity: 99.240 Percent Identity: 99.088

alignment_block:
 US-09-171-553B-6 x AAV09703 ..

Align seg 1/1 to: AAV09703 from: 1 to: 1974

1 MethisProThrLeuSerArgArgHisLeuProThrArgGlyGlyIupr 17
 |||||||
 1 ATGCATCCCAAGTTAACTGCGCCACCTCCGCACTCGGGGTGAGAGCC 50
 17 oluArgLeuAArgLleProLeuSerPheAlaSerLleAlaTrpPheLeu 34
 |||||||
 51 GAAAGACCTGAGAAATCCCTTAAGCTTGGCTGCCTCATGCGCTGGTCCCTTA 100
 34 hrLeuThrLleThrProGlnAlaSerSerLysArgLeuLeaSpSerSer 50
 |||||||
 101 CTGAACAATTAACCTCCAGGCGCAGTAGTAAGCCCTTATAGACAGCTCG 150
 51 AsnProHisArgProLeuSerLeuThrTrpLeuLleLeaSpProAsp 67
 |||||||
 151 AACCCCATAGACCTTATCCCTGACCTGCTGATTAAGCCCTGATAC 200
 67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84
 |||||||
 201 GGGTGTCACTGTAAATAGACACTGAGGTGTGCTGCTAGAGGACCTGGT 250
 84 rPrGlnLeuHisPheCysLeuAArgLeuLeaSpProAlaValLysSer 100
 |||||||
 251 GGCCTGAACATGCAATTTGCTGCTGCGATTTGATTAACCCGCGTTAAAGC 300
 101 ThrProProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117
 |||||||
 301 ACACCTCCCAACCTAGTCCGAGTTAGTGGTCTGATTTGCTGCGCAGGAC 350
 117 rGlyLeuGlnLysTyrCysGlyGlySerGlyGlySerPheCysArgArg 134
 |||||||
 351 AGAGAAAGACAAATATCTGGGGGTCTGGGGAATCTCTGTGAGGAGAT 400
 134 rPserCysValThrSerAsnAspGlyAspTrpLysTrpProLysLeu 150
 |||||||
 401 GGAGCTGCGTCACTCCACAGATGAGACAGTGAATGCGGATCTCTCTC 450
 151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyr 167
 |||||||
 451 CAGGACCGGGTAATAATCTCTTGTCAATTCGCGCCGCGGCAAGTCAAA 500
 167 smetMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAsp 184
 |||||||
 501 AATGATGAACATAATATAAGATAAGACTGCTCCCACTGACACTTAAT 550
 184 rLeuLysLleSerPheThrGlu..ArgLysThrGlyLysTyrSerLys 200
 |||||||
 551 ATCTAAAGATAGTTCACCTGAAAGAAAGAAACGAGAAATATTTCAAAG 600
 200 alaPlyStrPTrGlyLeuLeuLysAsnSerPheLeuLeuTyrGlyGly 216
 |||||||

601 TGCATTAATGATGAGCTGGCAATAGTTTTTATATAATATGCGGGG 650
 216 YAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGluP 233
 651 AGAGGGTCCACTTAAACATTCGCCCTAGATAGAGGGGAGCAAGAC 700
 233 rOPVaAlaMetGlyProAspLysValLeuAlaGluGlnGlyProPro 249
 701 CCCCTGGCAATGGGACCCGATGAAGTACTGGCTGACAGAGGGCCCCG 750
 250 AlAlaGluProProHisAsnLeuProValProGlnLeuThrSerLeu 266
 751 GCCCTGAGCCACCGCATTAAGTGGCGTCCCAATTAACCTCGCTGG 800
 266 gProAspIleThrGlnProProSerAsnSerThrGlyLeuIleProT 283
 801 GCCTGACATTAACAGCGCCCTAGCAAGTACCAGTGGATGATTCGA 850
 283 hrAsnThrProArgAsnSerProGlyValProValLysThrGlyLna 299
 851 CCAACAGCCTAGAAACTCCCGAGTGTCTGTTAAGACAGAGACAGA 900
 300 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrSpr 316
 901 CTCTCAAGTCTCAATCCAGGAGCCTTCCAGCCATCACTCCACGACCC 950
 316 oAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTyr 333
 951 TGATGCCACTTCTCTTGTGGCTTGTCTATCCCTCAGGGCCCTTATT 1000
 333 YrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHis 349
 1001 ATGAGGGATGCGCTAAAGAAATCAATGTGACCAAGACATAGA 1050
 350 AsnGluCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluVal 366
 1051 AATCAATGATACATGGGGTCCGAAATAAGCTTACCCTCACTGAAATTTTC 1100
 366 rGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeu 383
 1101 CGGGAAAGGGACATGCATAGGAAGTCCCCCATCCCAACACACTTT 1150
 383 YsTyrSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeuVal 399
 1151 GCTATAGTACTGCTTATAGACAGCCTCAGAAATCAGTATTAGTA 1200
 400 ProGlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCys 416
 1201 CCTGTTATACAGGTGGTGGCATCAATCTGGGTAAACCCCTGTGT 1250
 416 lSerThrSerValPheAsnGlnSerLysAspPheCysValMetValGln 433
 1251 TTCACCTGAGCTTCAACCAATCCAAAGATTGTGTGTCATGTCGTC 1300
 433 lValAlaProArgValTyrTyrHisProGluGluValValLeuAspGluTyr 449
 1301 TCTCTCCCGAGTGTACTACCATCTGAGGAAGTGTCTCTTATGAATAT 1350
 450 AspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeu 466
 1351 GACTATCGGTATACCGACCAAAAGAGAACCCGTATCCCTTACCTAGC 1400
 466 aValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAla 483
 1401 TGTAAATGCTCGATTTAGGGAGCGCCTTGGCGTAGAAGAGGACAGCTG 1450
 483 lAlaLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuGlyLueHis 499
 1451 CCTGATACAGGACCAAGAGCTAGAGAAAGACTGTGATGAGTACAT 1500
 500 AlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsn 516
 1501 GCGGCCATGACAGAAATCTCGAGCCTTAAAGAGTGTGTTAGCAACT 1550

516 uGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 533
 1551 AGAAGAGTCCCTGACTCTTGTGTGAAGTGTTCTTACAGAACCGAGG 1600
 533 lYLeuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaLeuLys 549
 1601 GATTAGATCTGCTGTTCTTAAGAGAGGTGGGTATGTGACGCCCTAAA 1650
 550 GluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSer 566
 1651 GAGAATGTGCTTATGTATGTACATCACTCAGAGCCATCAAGACTCAT 1700
 566 lAsnLysLeuArgLysLysLeuGluArgArgArgGluArgGluAla 583
 1701 GAACAAGCTTGAAGAAAGTTAGAGAGCGTCGAAGGAAAGAGGCGTG 1750
 583 sPGInGlyTyrPheGluGlyTyrPheAsnArgSerProTyrMetThr 599
 1751 ACCAGGGGTGGTTGAAGATGCTTCAACAGGTCTCTTGGATGACACC 1800
 600 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuTh 616
 1801 CTGCTTCTGCTCTGACGGGCCCTAGTACTCTCTCTCTTACTTAC 1850
 616 rValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArg 633
 1851 AGTTGGCGCTTGCTTAATTAAGTGTGTTGCTTGTAGAGACGAG 1900
 633 aISerAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeu 649
 1901 TGAGTGCAGTCCAGATCATGTACTTAGCAACAGTACCAAGGCCCTCTG 1950
 650 SerGlnGlyGluThrAspLeu 656
 1951 AGCCAGAGAAACTGACCTC 1971

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT.AAV82749

seq_documentation_block:
 ID AAV82749 standard; DNA: 3482 BP.

XX AAV82749;
 AC XX
 XX XX
 DT 25-FEB-1999 (first entry)
 XX XX
 DE Pig endogenous retrovirus (PERV) - B envelope (env) gene region.
 XX XX
 KW Pig endogenous retrovirus; PERV-A; envelope protein; PERV-B; subtype;
 KW probe; primer; detection; retrovirus; human tissue; xenotransplant;
 XX primary porcine tissue; human cell line; porcine cell line; ss.
 OS Pig endogenous retrovirus.
 XX XX
 FH Location/Qualifiers
 FT CDS 911..2884
 FT /*tag= a
 FT /product= envelope_protein

W09853104-A2.
 XX
 PD 26-NOV-1998.
 XX
 PE 18-MAY-1998; 98MO-GB01428.
 XX
 PR 16-MAY-1997; 97GB-0010154.
 XX
 PA (MED1-) MEDICAL RES COUNCIL.
 XX
 PI Sloye JP, Weiss RA;
 XX
 DR WPI, 1999-045324/04.
 DR P-PSDB; AAW85452.

XX Newly isolated nucleic acid probe capable of hybridizing to either
 PT the PERV-A or PERV-B env gene - useful in the detection of
 PT retroviruses, and their subtypes, in a sample of porcine/human
 PT tissue

PS Claim 3: Page 21-23: 36pp; English.

XX The present sequence encodes a pig endogenous retrovirus (PERV)-B
 CC envelope protein. PERV exists in two different subtypes, PERV-A and
 CC PERV-B. The differences are reflected in sequence divergence in the
 CC envelope genes. Probes and primers can be derived from the envelope
 CC (env) genes of PERV-A and PERV-B. The probes and primers are used in
 CC a method to detect retroviruses in a sample of porcine/human tissue,
 CC particularly primary porcine tissue and human cell lines that have been
 CC cultured in the presence of a porcine cell line, or human tissue from
 CC a patient with a xenotransplant. Subtype of PERV in a sample containing
 CC one of the PERV env genes can also be determined.

SQ Sequence 3482 BP; 927 A; 854 C; 867 G; 834 T; 0 other;

alignment_scores:

Quality: 3433.00 Length: 658
 Ratio: 5.257 Gaps: 2
 Percent Similarity: 99.240 Percent Identity: 98.176

alignment_block:
 US-09-171-553b-6 x AAV82749 ..

Align seg 1/1 to: AAV82749 from: 1 to: 3482

1 MethisProthrlleuSerArgrhlsleuprothrrargglyglupr 17
 911 ATGCATCCACAGTTAGCTGAGGCCACCTCCGAGCTCGGGTGGAGACC 960
 17 olysrargleuarqllleproleuSerPhealSerllealatrppheleut 34
 961 GAAAAGACGTGAGAAATCCCTTAAGCTTCGCCATCGCTGGTTCCTTA 1010
 34 hrleuThrllethrProglinaSerSerlysrargleuileaspserser 50
 1011 CTCTACATATACCTCCAGGCCAGTAAACGCTTATAGACAGCTCG 1060
 51 AsnProHlsArgProleuSerleuthrrpleuileaspproasprrh 67
 1061 AACCCCAATAGACCTTTATCCCTTACCTGCGTATATGACCCGATAC 1110
 67 rgllyalhrvalasnserrhrargglyvalalaproargglythrtrpt 84
 1111 GGGTGTCACTGTAATAGACACTCGAGGTGTGCTCTAGAGGACACTGT 1160
 84 rpprogluleuHlsPheCysleuargleuileasproalvalyllyser 100
 1161 GGCCTGAAGCTGATTTGCTCGCTCGATTGATTAACCCGCTGTAAAGC 1210
 101 ThrProPasnleuValArgSerlyrlyphethrCysCysprogllyth 117
 1211 ACACCTCCCAACCTAGTCCGATGATTAGGTTCTATTGCTGCCAGGAC 1260
 117 rgllyslulstlyrCysglyserglygluSerPheCysargrgrt 134
 1261 AGAGAAAGAGAAATACGTGTGGGGTCTGGGGAATCCTTCTGTGGAGAT 1310
 134 rpsrCysValThrSerasnspgllyasprtplysttpproilserleu 150
 1311 GGAAGTGGTCACTCCACAGATGAGAGCTGGAATGCGCATCTCTCTC 1360
 151 GlaspargVallysrPheSerPheValasnserrglyprogllylytyly 167
 1361 CAGGACGGGTAAATTCCTCTTGTCAATTCGCGCCCGGAGCACTACAA 1410
 167 SmetleuThlyleuThlysrlysrCysSerProSerAspleuaspr 184

1411 AGTATGAACTATATATAAGTAAAGAGCTGCTCCCATACACTTACATTT 1460
 184 yrleuylsIleSerPheThrGlu..ArglyThrGlyLysTySerlysv 200
 1461 ATCTAAAGATTAAGTTTCACTGAAAAGAGAAAACAGAAAATATTCAAAAG 1510
 200 alaaplysrtrpThrGluLeuGlyasnSerPheleu..LeuTyrglyglyl 216
 1511 TGGATAAATGGTATGAGCTGGGGAATAGTTTATTAATATGCGGGGG 1560
 216 yalaglySerThrleuthrIleArgleuArgIleGlyThrGlyThrGlu 233
 1561 AGCAGGGTCCACTTAACCATTCGCCCTTAGATAGAGAGCGGGGACAGAC 1610
 233 ropovalalameGlyProaspIysValleualagluglmglypro 249
 1611 CCCTGTGGGAGTGGGACCCGATAAAGTACTGGCTGAACAGGGGCCCCCG 1660
 250 AlaLeuGluProProHlsAsnleuProvalproglleuthrSerleuar 266
 1661 GCCCTGGAGCCAGCCATACCTGCGGTGCCCAATTAACCTGCTGGC 1710
 266 gProAspIleThrGlnProProSerAsnSerThrThrGlyLeuilePro 283
 1711 GCCTGACATATACACAGCGCGCTACCAACGATACACACTGATGATTCCTA 1760
 283 hrAsnThrProArgAsnSerProglyValProvallysrhrglyGlnarg 299
 1761 CCAACAGCGCTAGAAACTCCCGAGGTGTCTCTGTTAAGACAGACAGAGA 1810
 300 leuPheSerleuileGlnGlyAlaPheGlnAlaIleasnserrhraspr 316
 1811 CTCTTCAGTCTCATCGAGGAGCTTCCAAAGCATCAATCCACCGGACC 1860
 316 oAspAlaThrSerSerCysTrpleuCysleuSerSerGlyProProtyr 333
 1861 TGAAGCCACTTCTCTGTTGGCTTGTCTATCTCTCAAGGCGCCCTTATT 1910
 333 yrglyGlyMetAlaLysGlnArgLysPheasnValThrLysGluHlsarg 349
 1911 ATGAGGGGATGGCTAAAGAGAAATTCATATGACCAAAAGAGCATAGA 1960
 350 AsnGlnCysThrTrpGlySerArgAsnLysleuThrleuthrGlnvalse 366
 1961 AATCATATGTCATGGGGGCTCCGAAATAGCTTACCTCACTGAAAGTTTC 2010
 366 rgllyslglyThrCysIleGlyLysAlaProProSerHlsGlnHlsleuc 383
 2011 CGGGAAGGGACATGCATAGGAAAGCTCCCGCATCCACACACACTTT 2060
 383 ysrlysrThrValValThrGluGlnAlaSerGluasnGlnTyrlleuVal 399
 2061 GCTATAGTACTGTGTTATGAGCAGGCCCTCAAGAAATCAGTATTAGTA 2110
 400 ProGlyTyraAsnArgTrpTrpAlaCysAsnThrGlyleuthrProCysva 416
 2111 CTTGTTATACAGGTGGTGGCATGCAATACGGGTTAACCCCTGTGT 2160
 416 lSerThrSerValPheasnGlnSerlysrPheCysValIleValGlnI 433
 2161 TTCACACTGAGTCTTCACCAATCCAAAGATTTCTGTGCATAGTGTCCAA 2210
 433 lvalAlaProargValTyrrhIsProgluGlnValleuasprglutyr 449
 2211 TCGTCCCGCCAGTGTACTACATCTGAGAGATGTGCTTGTGATTAAT 2260
 450 AspTyrrArgTyraAsnArgProlysrArgGluProvalSerleuthrleu 466
 2261 GACTATCGGTATACCGACCAAAAGAAAGAACCCGTATCCCTTACCTTAC 2310
 466 aValIleleuGlyleuGlyThrAlaValalaglyThrGlyThrAla 483


```

131 yArGArgrtPpSerCysValThrSerAsnAspGlyAspTrpLysTrpPro 147
11:.....
611 GCAGCAAGAGGAGCTGCATACCTCTTAATGATGGATGGAATGGAATGGCCA 660
148 lIeSerLeuGlnAspArgValLysPheSerPheValAsnSer..... 161
661 GTCCTCAGCAAGACAGAGTAAGTACTCTTTGTTAACATCATCCACG 710
162 .....GlyProGlyLysTrpLys.....MetMetL 170
711 TTATAATCAATTTAATTAATGATGGATGGATGGAATGGCAACG 760
170 yLeuThrLysAspLys.....SerCysSerProSerAsp 181
761 GGGTACAAAAGATGTACGAATAAAGCAAAATGATCATCTGTTAGAC 810
182 LeuAspTrpLeuLyslIeSerPheThrGlu..ArgLysThrGlyLysTrp 197
811 CTAGATTACTTAATAAATAAGTTCCTGCTGAAAAAGAAAAACAAGAAAT 860
198 SerLysValAspLysTrpTrpGlyLeuGlyAsnSer..PheLeuLeuTrp 214
861 TCAAAAAGGGGTAAATGTATATCTTGGGAAATAGTATAC.....TATG 904
214 lYgLYgLYAlA.....GlySerThrLeuThrIleArgLeuArgIle 227
905 GAGGCTCTGGGAGAAAGAGATCTGTTGCTGACTATTCGCTCAGATA 954
228 GluThrGlyThrGluProProValAlaMetGlyProAspLysValLeuAl 244
955 GAACATCGAGTGGAACTCCGGTGTCTATAGGACCAATTAAGGGTTTGGC 1004
244 aGluGlnGlyProProAlaLeuGluProProHisAsnLeuProValPro 261
1005 CCAACAGAGACCTCCATCCAAAGACAG..... 1032
261 lInLeuThrSerLeuArgProAspIleThrGlnProProSerAsnSerThr 277
1033 .....AGGCCATCTCCTCAACCCCTGATTACATAATCAAC 1068
278 ThrGlyLeuIleProThrAsnThrProArgAsnSerProGlyValAlaPro 294
1069 TCTGGATAGTCCCACT.....GAGCCTAACATCCTACTAT 1103
294 lLysThrGlyLysArgLeuPheSerLeuIleGlnGlyAlaPheGlnAla 311
1104 TAAACAGAGGGGAAACTTTTAGCTCATCCAGAGGAGCTTTCAAGCTC 1153
311 lAsnSerThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSer 327
1154 TTAACTCCAGCACTCCAGAGGCTACTCTTCTGTTGGCTTGGCTTAGCT 1203
328 SerGlyProProTrpTrpGlyGlyMetAlaLysGlyLysArgLysPheAsn 344
1204 TCGGGCCCACTTACTATGAGGAAATGGCTAGAGGAGGAAATTCATGT 1253
344 lThrLysGlnHisArgAsnGlnCysThrTrpGlySerArgAsnLysLeu 361
1254 GACAAAGGACATAGAGCAATGTACATGGGATCCCAAAATATAGCTTA 1303
361 hrlLeuThrGluValSerGlyLysGlyThrCyslIeGlyLysAlaProPro 377
1304 CCTTACAGAGTTCTGGAAAAAGCACTGCATAGGAAAGTGTCCCA 1353
378 SerHisGlnHisLeuCysTrpSerThrValValTrpGlnGlnAlaSerG 394
1354 TCCGACCAACACCTTTGTAAACACACGAGAGCTTTATATGCAACCTGTA 1403
394 uAsnGlnTrpLeuValProGlyTrpAsnArgTrpTrpAlaCysAsnThrG 411
1404 GAGTCATATATCTGCTACCTGTTATGACAGGTGTGGGCAATGTAACTG 1453
411 lYLeuThrProCysValSerThrSerValPheAsnGlnSerLysAspPhe 427

```

```

1454 GATTAAACCCCTGTGTGTTCCACCTGTGTTTCAACCAAACTAAAGACTTT 1503
428 CysValMetValGlnIleValProArgValLysTrpHisProGlnGluVal 444
1504 TGCCTTATGTGCCAAATTTGTCCTCCCGGTTGACTACTATCCGAAAAAGC 1553
444 lValLeuAspGlnTrpAspTrpArgTrpAsnArgProLysArgLysPro 461
1554 AGTCCTTGATGATATATGACTATACATATATCGGCCAAAGAAAGAGCCCA 1603
461 aLSerLeuThrLeuAlaValMetLeuGlyLeuGlyThrAlaValAlaGly 477
1604 TATCCCTGACACCTFAGCTGTAATGCTCGGATGGGAGTGGCTGCAGGGGTG 1653
478 GlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeuGlyLysG 494
1654 GGAACAGGAACGGCTGCCCTAATCAGAGACCCGCAACAGCTGGAGAAAG 1703
494 YLeuGlyGlnLeuHisAlaAlaMetThrGluAspLeuArgAlaLeuLysG 511
1704 ACTTAGTAACCTACATCGAATTGTACGGAAGATCTCCAAAGCCCTAGAAA 1753
511 lUsrValSerAsnLeuGlnGlnSerLeuThrSerLeuSerGlyValAla 527
1754 AATCTGCAGTAACCTGAGGAAATCCCTAACCTCTTATCTGAAGTGT 1803
528 LeuGlnAsnArgArgGlyLeuAspLeuLeuPheLeuArgGlyLysGly 544
1804 CTACAGAACAGAAAGGGGTAGATGTGTATTTCTAAAGAGAGGGT 1853
544 uCysAlaAlaLeuLysGlnGlyLysCysPheTrpValAspHisSerGly 561
1854 ATGTGTAGCTTTAAAGAGCAATGCTCTTCTATGTAATGATCAGCTAGAG 1903
561 lAlIeArgAspSerMetAsnLysLeuArgLysLysLeuGlnArgArg 577
1904 CCATCAGACAGACTCCATGAGCAAGCTTAGAGAAAGTTAGAGAGCGCTGCA 1953
578 ArgGlnArgGlnAlaAspGlnGlyTrpPheGlnGlyTrpPheAsnArg 594
1954 AGGGAAGAGAGGCTGCACAGGGGTGTGTAAGAGATGCTTCAACAGGTC 2003
594 rProTrpMetThrThrLeuLeuSerAlaLeuThrGlyProLeuValVal 611
2004 TCCCTGGATACCACTGCTTCTGCTCTGACGGGGGCCCTTAGTAGTGC 2053
611 euleuLeuLeuLeuThrValGlyProCysLeuIleAsnArgPheValAla 627
2054 TGCCTCGTTACTTACAGTTGGGCTTGCCTTATATAGGCTTGTGTGCC 2103
628 PheValArgGlnArgValSerAlaValGlnIleMetValLeuArgGlnG 644
2104 TTTGTTAGAGAACAGAGTGAAGTGCATGCCAATCATGTGTTAGGCACAA 2153
644 nTYrGlnGlyLeuLeuSerGlnGlyLysTrpAspLeu 656
2154 GTACCAAGGCTTCTCAGCCAAAGAGAAACTGACCTTC 2190

```

seq_name: /SID52/gcgdata/geneseq/geneseqn/NA2001.DAT: AAC67022

seq_documentation_block:

ID AAC67022 standard; DNA: 4918 BP.

XX AAC67022;

XX 27-MAR-2001 (first entry)

XX PERV env protein coding sequence SEQ ID NO: 22.

XX xenotransplantation; infectious agent; vaccine; ds.

XX Porcine endogenous retrovirus.

XX WO200071726-A1.
 XX 30-NOV-2000.
 XX 24-MAY-2000; 2000MO-US14296.
 XX 24-MAY-1999; 9905-0135631.
 XX (MAYO-) MAYO MEDICAL VENTURES.
 XX Federspiel MJ;
 XX WPI; 2001-032041/04.
 XX Inhibiting or preventing infectious agent transmission in mammalian
 PT transplant recipients, by introducing recombinant DNA comprising DNA
 PT encoding extracellular proteins of the agent into donor cells, such as
 PT swine cells -
 XX
 PS Claim 16; Page 109-111; 144pp; English.
 XX
 CC The present invention provides a method to prevent the transmission of
 CC infectious agents during xenotransplantation. This involves introducing
 CC to donor swine cells a recombinant DNA encoding a peptide fragment from
 CC the infectious agent, and then introducing these cells into the
 CC transplant recipient.
 XX
 SQ Sequence 4918 BP; 1382 A; 1173 C; 1183 G; 1180 T; 0 other;

alignment_scores:
 Quality: 2509.50 Length: 670
 Ratio: 4.357 Gaps: 11
 Percent Similarity: 85.970 Percent Identity: 70.896

alignment_block:
 US-09-171-553b-6 x AAC67022 ..

Align seg 1/1 to: AAC67022 from: 1 to: 4918

1 MethSerProThrLeuSerArgArgHisLeuProThrArgGlyGlyLeuPr 17
 ||||||||||||||||||||||||||||||||||||||||||||||||
 2187 ATGCATCCACGCTTAAGCCGGCCACCTCCCGATTCGGGGTGGAAACCC 2236
 ||||||||||||||||||||||||||||||||||||||||||||||||
 17 OlyshArgLeuArgLeuProLeuSerPheAlaSerIleAlaTrpPheLeu 34
 ||||||||||||||||||||||||||||||||||||||||||||||||
 2237 GAAAGAGCTGAAGAACCCCTTAAGCTTCGCTCCATCGCGTGTCTCTTA 2286
 ||||||||||||||||||||||||||||||||||||||||||||||||
 34 hLeuThrIleThrProGlnAlaSerSerLySArgLeuIleAspSerSer 50
 ||||||||||||||||||||||||||||||||||||||||||||||||
 2287 CTCTGTCAATACCTCCTCAAGTTAATGCTAAACGCCCTGTGTGACAGCCG 2336
 ||||||||||||||||||||||||||||||||||||||||||||||||
 51 AsnProHisArgProLeuSerLeuThrTrpLeuIleIleAspProAspPth 67
 ||||||||||||||||||||||||||||||||||||||||||||||||
 2337 AACCTCCATTAACCTTATCTCTCACCTGGTTACTTACTGACTCCGGTAC 2386
 ||||||||||||||||||||||||||||||||||||||||||||||||
 67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84
 ||||||||||||||||||||||||||||||||||||||||||||||||
 2387 AGGTATTAATATTACAGACACCAAGGGAGGCTCCCTGGGAGACCTGTGT 2436
 ||||||||||||||||||||||||||||||||||||||||||||||||
 84 rProGlnLeuHisPheCysLeuArgLeuIleAsnProAlaVal..... 98
 ||||||||||||||||||||||||||||||||||||||||||||||||
 2437 GCCCTGAATTAATGTCTGCTGCATCAGTAATCCCTGGTCTCAATGAC 2486
 ||||||||||||||||||||||||||||||||||||||||||||||||
 99 LysSerThrProProAsnLeuValArgSerTrpGlyPheTrpCysAsp 115
 ||||||||||||||||||||||||||||||||||||||||||||||||
 2487 CAGGCCACACCCCGCATGCTACTCGTCTTAAGGGTTTACGTTTGCC 2536
 ||||||||||||||||||||||||||||||||||||||||||||||||
 115 oGly...ThrGlnGlyGlyTrpCysGlyGlySerGlyLeuSerPheC 131
 ||||||||||||||||||||||||||||||||||||||||||||||||
 2537 AGGACCCCAATATATGAAGAAATATTGTGAATCTCAGGATTTCTTTT 2586

131 ysArgArgTrpSerCysValThrSerAsnAspGlyAspTrpLysTrpPro 147
 ||||||||||||||||||||||||||||||||||||||||||||||||
 2587 GCAAGCAATGAGCTCGTAACTTTAAATGATGGAATGGAATGGCCA 2636
 ||||||||||||||||||||||||||||||||||||||||||||||||
 148 IleSerLeuGlnAspArgValLysPheSerPheValAsnSer..... 161
 ||||||||||||||||||||||||||||||||||||||||||||||||
 2637 GTCTCTCAGCAAGACAGAGTAAGTACTCTTTGTTACAAATCTTACGAC 2686
 ||||||||||||||||||||||||||||||||||||||||||||||||
 162GlyProGlyLysTrpLys.....MetMetL 170
 ||||||||||||||||||||||||||||||||||||||||||||||||
 2687 TTATTAATCAATTAATTATGGCCATGGAGATGGAATGGCAACACAC 2736
 ||||||||||||||||||||||||||||||||||||||||||||||||
 2737 GGGTACAAAAGATGTACGAAATTAACCAATAATAGCTGCTATTCGTTAGC 2786
 ||||||||||||||||||||||||||||||||||||||||||||||||
 182 LeuAspTrpLeuLysIleSerPheThrGln..ArgLysThrGlyLysTrp 197
 ||||||||||||||||||||||||||||||||||||||||||||||||
 2787 CTAGATTACTTAAATAATTAAGTTTCACTGAAAAGAAAACAGAAATAT 2836
 ||||||||||||||||||||||||||||||||||||||||||||||||
 198 SerLysValAspLysTrpTrpGlnLeuGlyAsnSer..PheLeuLeuTrpG 214
 ||||||||||||||||||||||||||||||||||||||||||||||||
 2837 TCAAAAGTGGGTAAATGATGTATCTTGGGGAATAGTGTAC.....TATG 2880
 ||||||||||||||||||||||||||||||||||||||||||||||||
 214 LysGlyGlyAla.....GlySerThrLeuThrIleArgLeuArgGly 227
 ||||||||||||||||||||||||||||||||||||||||||||||||
 2881 GAGGCTCTGGGAGAAAGAAAGATCTGCTGTGACTATTCGCTTCAAGAAA 2930
 ||||||||||||||||||||||||||||||||||||||||||||||||
 228 GlnThrGlyThrGlnProProValAlaMetGlyProAspLysValLeuAl 244
 ||||||||||||||||||||||||||||||||||||||||||||||||
 2931 GAAACTCAGATGGAACCTCGGTTGCTATAGACCAATAGAGGTTTGGC 2980
 ||||||||||||||||||||||||||||||||||||||||||||||||
 244 agLingLysProProAlaLeuGlnProProHisAsnLeuProValProG 261
 ||||||||||||||||||||||||||||||||||||||||||||||||
 2981 CGAAACAAGGACCTCCCAATCCAAAGACAG..... 3008
 ||||||||||||||||||||||||||||||||||||||||||||||||
 261 hLeuThrSerLeuArgProAspIleThrGlnProProSerAsnSerThr 277
 ||||||||||||||||||||||||||||||||||||||||||||||||
 3009AGGCCATCTCTTACCCCTCTGATTAACATTAACACACC 3044
 ||||||||||||||||||||||||||||||||||||||||||||||||
 278 ThrGlyLeuIleProThrAsnThrProArgAsnSerProGlyValProVa 294
 ||||||||||||||||||||||||||||||||||||||||||||||||
 3045 TCTGATCAGATCCCACT.....GACCTTAACATCACTAT 3079
 ||||||||||||||||||||||||||||||||||||||||||||||||
 294 LysThrGlyGlnArgLeuPheSerLeuIleGlnGlyAlaPheGlnAla 311
 ||||||||||||||||||||||||||||||||||||||||||||||||
 3080 TAAACAGGGGGCGAAACTTTTAACTCAATCCAGGAGCTTTTCAAGCTC 3129
 ||||||||||||||||||||||||||||||||||||||||||||||||
 311 LeuAsnSerThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSer 327
 ||||||||||||||||||||||||||||||||||||||||||||||||
 3130 TTAACTCCAGGACTCCAGAGGCTACCTCTTGTGTGGCTTGTAGCT 3179
 ||||||||||||||||||||||||||||||||||||||||||||||||
 3180 TCGGGCCACCACTTACTATAGAGGAATGGCAGAGGAGGAATTAATCAT 3229
 ||||||||||||||||||||||||||||||||||||||||||||||||
 344 LThrLysGlnHisArgAsnGlnCysThrTrpGlySerAlaGlnLysLeu 361
 ||||||||||||||||||||||||||||||||||||||||||||||||
 3230 GACAAAGAACATAGAGACCAATGATGAGGATCCCAAAATAATCACTTA 3279
 ||||||||||||||||||||||||||||||||||||||||||||||||
 361 hLeuThrGlnValSerGlyLysGlyThrCysTrpIleGlyLysAlaProPro 377
 ||||||||||||||||||||||||||||||||||||||||||||||||
 3280 CCTTACTAGGTTTCTGGAAAGGACCTGCTATAGGATGGTGTCCCA 3329
 ||||||||||||||||||||||||||||||||||||||||||||||||
 378 SerHisGlnHisLeuCysTrpSerThrValValTrpGlnGlnAlaSerG 394
 ||||||||||||||||||||||||||||||||||||||||||||||||
 3330 TCCACACCAACCTTTGATACACACAGTGAAGCCTTAATGCAACCTCTG 3379
 ||||||||||||||||||||||||||||||||||||||||||||||||
 394 uAsnGlnTrpLeuValProGlyTrpAsnArgTrpTrpAlaCysAsnTrpG 411
 ||||||||||||||||||||||||||||||||||||||||||||||||
 3380 GAGTCAATATCTGTACTGTGTATGACAGCTGTGCGCATGTAAATACG 3429
 ||||||||||||||||||||||||||||||||||||||||||||||||
 411 LysLeuThrProCysValSerThrSerValPheAsnGlnSerLysAspPhe 427


```

|||||
3430 GATTAAACCCCTGTGTTCCACCTTGTTTCCACCAACTAAGACTTT 3479
428 CysValMetValGlnIleValProArgValTyrTyrHisProGluGluVal 444
3480 TGGGTATGTTCCAAATGTGCCCCGGGTGTTACTATATCCGGAAGAAC 3529
444 ValLeuAspGluTyrAspTyrArgTyrAsnArgProGlySerGluProVal 461
3530 AGTCCTTGATGATATGACTATGATATATGCGCCCAAAAAGAGACCCA 3579
461 aIseIleuThrIleuAlaIleuMetLeuGlyLeuGlyThrAlaValGlyVal 477
3580 TATCCCTGACTAGCTGTAAGCTCGATGAGAGTGGAGTGGCTGAGGCGTG 3629
478 GlyThrGlyThrAlaIleuIleThrGlyProGlnGlnLeuGluGluVal 494
3630 GGAACAGAAAGCGCTGCCCTAATCACAGGACCCGACAGCTGGAGAAAG 3679
494 yLeuGlyGluLeuHisAlaIleuMetThrGluAspLeuArgAlaLeuLysG 511
3680 ACTTAGTAACCTCATCATGATGTAAGGAGATCTCCAAAGCCCTAGAAA 3729
511 IuSerValSerAsnLeuGluGluSerLeuThrSerLeuSerGluValVal 527
3730 AATCTGCACTAAGCTGAGAGATCCCTAACCCTTATCTGAAGTGTGTT 3779
528 LeuGlnAsnArgArgGlyLeuAspLeuPheLeuArgGluGlyLys 544
3780 CTACAGACAGAAAGGGGTAGATCTGTATTCTTAAAGAGAGAGATT 3829
544 uCysAlaIleuLeuGluGluGlyCysPheTyrValAspHisSerGlyVal 561
3830 ATGTGTAGCTTGAAGAGAAATGCTGTTTATGTGATCATTCAGGCG 3879
561 IaIleuArgAspSerMetAsnLysLeuArgLysLysLeuGluArgArgArg 577
3880 CCATCAGAGACTCCATGACAAAGCTTAGAAAGGTTGGAGAGCGCTCA 3929
578 ArgGluArgGluAlaAspGlnGlyTyrPheGluGlyTyrPheAsnArgSe 594
3930 AGGGAAGAAAGAACTACTCAAGGTGTTGAGGATGTTCAACAGGTC 3979
594 rProTyrMetThrThrIleuLeuSerAlaLeuThrGlyProLeuValVal 611
3980 TCCTTGTTGGTACCTACTTCTGCTTAAACAGGACCTTAATAGTCC 4029
611 euLeuLeuLeuLeuThrValGlyProCysLeuIleAsnArgPheValAla 627
4030 TCCCTCCGTACTACAGAGTTGGCCAGTATTTATTAACAAGTTAATTGCC 4079
628 PheValArgGluArgValSerAlaValGlnIleMetValLeuArgGlnG 644
4080 TTCATGTAGAGAGAAATAGTGAAGTCCAGATCATGTACTTAGACAACA 4129
644 nTyrGln 646
4130 GTACCAA 4136
seq_name: /SID2/gcgdata/geneseq/geneseqn/NA2001.DAT: AAC67018
seq_documentation_block:
ID AAC67018 standard; DNA; 1980 BP.
XX
XX AAC67018;
AC
XX
XX 27-MAR-2001 (first entry)
DT
XX
XX PERV-1-15 env protein coding sequence SEQ ID NO: 18.
DE
XX
XX Xenotransplantation; infectious agent; vaccine; ds.
XX
OS Porcine endogenous retrovirus.

```

```

XX
XX W0200071726-A1.
XX
XX 30-NOV-2000.
XX
XX 24-MAY-2000; 2000WO-US14296.
XX
XX 24-MAY-1999; 99US-0135631.
XX
XX (MAYO-) MAYO MEDICAL VENTURES.
XX
XX Federspiel MJ;
XX
XX WPI: 2001-032041/04.
XX
XX Inhibiting or preventing infectious agent transmission in mammalian
XX transplant recipients, by introducing recombinant DNA comprising DNA
XX encoding extracellular proteins of the agent into donor cells, such as
XX swine cells -
XX
XX PS Claim 16; Page 100-101; 144pp; English.
XX
XX CC The present invention provides a method to prevent the transmission of
XX CC infectious agents during xenotransplantation. This involves introducing
XX CC to donor swine cells a recombinant DNA encoding a peptide fragment from
XX CC the infectious agent, and then introducing these cells into the
XX CC transplant recipient.
XX
XX SQ Sequence 1980 BP; 582 A; 452 C; 447 G; 499 T; 0 other;

```

```

alignment_scores:
  Quality: 2497.50      Length: 669
  Ratio: 4.343          Gaps: 11
  Percent Similarity: 85.949   Percent Identity: 70.852
alignment_block:
US-09-171-553B-6 x AAC67018 ..

```

Align seg 1/1 to: AAC67018 from: 1 to: 1980

```

1 MethIspProThrLeuSerArgArgHisLeuProThrArgGlyGluPr 17
|||||
1 ATGCATCCACGCTTAAGCGGGGCCACCTCCCGATTGCGGGTGAAGACC 50
17 oLysArgLeuArgGlyLeuProLeuSerPheAlaSerIleAlaTrpPheLeu 34
|||||
51 GAAAGAGACTGAAATCCCTTAAGCTTGCCCTCATCGCTGTCCTTA 100
34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50
|||||
101 CTCTGTCAATTAAGCTCTCAAGTAAATGTAAGCGCTGTGACAGCCG 150
51 AsnProHisArgProLeuSerLeuThrTyrIleIleIleAspProAspTh 67
|||||
151 AACTCCCATTAACCCCTTATCTCTACCTGTTACTTACTGACTCCGCTTA 200
67 rGlyValIleThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84
|||||
201 AGGTATTAATATTAAACAGCAGCTCAAGGAGGAGCTCCCTTGGGAGCTG 250
84 rPProGluLeuHisPheCysLeuArgLeuIleAsnProAlaVal..... 98
|||||
251 GGCCTGAATTAATATGCTGCTTCATGATGATATGCTGCTGCTGCTG 300
99 LysSerThrProProAsnLeuValArgSerTyrGlyPheTyrCysCysPr 115
|||||
301 CAGGCCACACCCCGGATGACTCCGCTGCTGAGGATTTTACGTTGGCC 350
115 oGly...ThrGluysGluLysTyrCysGlyGlySerGlyLysSerPheC 131
|||||
351 AGGCCCCCAATTAATGAAGAAATATTGTGGAATCTCTCAGGATTTCTTT 400

```

```

131 ysArgArgTrpSerCysValThrSerAsnAspGlyAspTrpLysTrpPro 147
1401 GCAGCAAGATGAGCTGCGTAACCTTCAATGATGGAGATGGAAATGGCCA 450
148 IleSerLeuGlnAspArgValLysPheSerPheValAsnSer..... 161
451 GTCTCTCAGCAAGACAGAGTAAAGTACTCTTTGTTAACATCACTCCACAG 500
162 .....GlyProGlyLysTyrLys.....MetMetL 170
501 TTATATCAATTTAATTATGGCCATGGAGATGGAAAGATGGCAACAC 550
170 yslLeuTyrLysAspLys.....SerCysSerProSerAsp 181
551 GGGTCAAAAAGATGTACCAATAATAGCAATAAGCTGATCTGTTAGAC 600
182 LeuAspTyrLeuLysLysLeuSerPheThrGlu..ArgLysThrGlyLysTyr 197
601 CTGATTTACTTAAATAAGTTTCACTCAAAAAGAAACAAAGAAATAT 650
198 SerLysValAspLysTrpTyrGluLeuGlyAsnSerPheLeu..TyrG 214
651 TCAAAAAGTGGTAATGATGATGTTGGGAATAGT...GTACTATAGAG 697
214 IyGlyGly.....AlaGlySerThrLeuThrIleArgLeuArgIleGlu 228
698 GCTTGGGAGAAAGAAAGAGATCTGTTGACATTCGCTCAGATTAAGA 747
229 ThrGlyThrGluProProValAlaMetGlyProAspLysValLeuAlaG 245
748 ACTCAGATGAGACCTCCGCTGCTATAGACCAATAAGGGTTGGCCGA 797
245 uGlnGlyProProAlaLeuGluProProHisAsnLeuProValProGlnL 262
798 ACAAGGACTCCATCCAGAACAG..... 822
262 eutHrSerLeuArgProAspIleThrGlnProProSerAsnSerThrThr 278
823 .....AGCCATCTCCCTTAAACCCCTGATTAACATACAAACCTCT 861
279 GlyLeuIleProThrAsnThrProArgAsnSerProGlyValProValL 295
862 GGATCAGTCCCCACT.....GAGCTTACATCACTATATTA 896
295 sThrGlyGlnArgLeuPheSerLeuIleGlnGlyAlaPheGlnAlaIle 312
897 AACAGGGCGGAACCTTTTAACCTCAATCCAGGAGCTTTCAACCTCTTA 946
312 snSerThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSerSer 328
947 ACTCCACGACTCCAGAGGCTACCTCTTGTGCTTGGCTTACCTTCG 996
329 GlyProProTyrTyrGluGlyMetAlaLysGluArgLysPheAsnValTh 345
997 GGGCCACCTTAATATGAGGATGGCTAGAGAGGAAATTCATATGTGAC 1046
345 rLysGlnHisArgAsnGlnCysThrTrpGlySerArgAsnLysLeuThr 362
1047 AAAAGAACATAGAGACCAATGTACATGGGATCCCAAAATTAAGCTTACC 1096
362 eutHrGluValSerGlyLysGlyThrCysIleGlyLysAlaProProSer 378
1097 TTACTGAGGTTTCTGAAAAGGCACTGCATAGGAGATGTTCCGCCATCC 1146
379 HisGlnHisLeuCysTyrSerThrValValTyrGluGlnAlaSerGluAs 395
1147 CACCAACACCTTGTATACACACATGAGAGCTTAATAGAACCTCTGAGAG 1196
395 nGlnTyrLeuValProGlyTyrAsnArgTPTTPAlaCysAsnThrGlyL 412
1197 TCAGTATCTGTACTGCTGTTATGACAGGTGGTGGCATGTAATACCTGAGAT 1246
412 eutHrProCysValSerThrSerValPheAsnGlnSerLysAspPheCys 428

```

```

1247 TAAACCTTGTTTCCACCTTGTTTCCAAACAAACAAAGCTTTTGC 1296
429 ValMetValGlnIleValAlaProArgValTyrTyrHisProGluGluVal 445
1297 GTTATGGTCCAAATTTGTCCTCCCGGGGTGACTACTATCCGAAAAAGCACT 1346
445 IleuAspGluTyrAspTyrArgTyrAsnArgProLysArgGluProValS 462
1347 CCTGATGATATGATGATATGATATTAATCGGCAAAAAAGAGAGCCATAT 1396
462 erLeuThrLeuAlaValMetLeuGlyLeuGlyThrAlaValAlaGlyVal 478
1397 CCTGACACTGACTGTAATGCTCGGATTTGGAGTGGCTGACGGCTGGGA 1446
479 ThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeuGlyLysGly 495
1447 ACAGGAACGGCTGCCCTTAATCAGAGACCCGCAACAGCTGGAGAAAGACT 1496
495 uGlyGluLeuHisAlaAlaMetThrGluAspLeuArgAlaLeuLysGluS 512
1497 TAGTAACCTACATCAATTAATTAACGAAATTCCTCAAGCCCTAGAAAT 1546
512 erValSerAsnLeuGluGluSerLeuThrSerLeuSerGluValValLeu 528
1547 CTGACAGTAACCTGAGAGAAATCCCTAACCTCTTATCTGAAAGGGTTCTA 1596
529 GlnAsnArgArgGlyLeuAspLeuPheLeuArgGluGlyLysLeu 545
1597 CAGAGCAAGAGGGGTAGATCTGTTATTTCTAAAGAAAGAGAGATTATG 1646
545 sAlaAlaLeuLysGluGluCysCysPheTyrValAspHisSerGlyAla 562
1647 TGTAACCTTAAGAGAGAAATGCTGTTTATGTGATCATTCAGGGGCCA 1696
562 leuArgAspSerMetAsnLysLeuArgLysLysLeuGluArgAlaGArg 578
1697 TCAGAGCTCCTCAATGAAACAGCTTGAAGAAAGTTGAGAGAGCTCGAAG 1746
579 GluArgGluAlaAspGlnGlyTrpPheGluGlyTrpPheAsnArgSerPr 595
1747 GAAAGGAACACTACTCAAGGCTGTTTGAAGGATGTTTAAACAGCTTCC 1796
595 oTrpMetThrThrLeuLeuSerAlaLeuThrGlyProLeuValValLeu 612
1797 TTGTTGGCTACCTCTTCTCTTAAACAGAGCCCTTAATAGCTTCC 1846
612 eutLeuLeuThrValGlyProCysLeuIleAsnArgPheValAlaPhe 628
1847 TCTGTACTCACAGTTGGGCCATGTATTAATTAACAAATTAATGCTTTC 1896
629 ValArgGluArgValSerAlaValGlnIleMetValLeuArgGlnGlnT 645
1897 ATTAGAGAAAGAAATAGTGCAGTCCAGATCAATGATCACTTAAGCAACAGTA 1946
645 rGln 646
1947 CCAA 1950

```

seq_name: /SIDS2/gcdata/geneseq/geneseq/NA2001.DAT:AAF77726

seq_documentation_block: ID AAF77726 standard; DNA: 7333 BP.

AAF77726;

23-MAY-2001 (first entry)

Defective retroviral genome isolated from PK-15 cell line.

Retrovirus; graft transplantation; xenotransplantation; PK-15 cell line;

ss.

OS Unidentified.
 XX US6190861-B1.
 PN 20-FEB-2001.
 PD 13-DEC-1996: 9605-0766528.
 PF 14-DEC-1995: 9505-0572645.
 PR (GEHO) GEN HOSPITAL CORP.
 XX Fishman JA;
 XX WPI: 2001-256211/26.
 DR P-P-SDB: AAB73282, AAB73283, AAB73284.
 XX Assessing risk of endogenous retroviruses in clinical practice and in
 PT xenotransplantation, comprises using probe sequences derived from swine
 PT or miniature swine retroviral genome -
 XX Claim 1; Fig 2; 127pp: English.
 XX The present invention relates to a method for screening a cell or tissue
 CC for the presence or expression of a retrovirus (RV), comprising
 CC contacting a target nucleic acid from the cell or tissue with a second
 CC nucleic acid from the present invention (e.g. the present sequence or a
 CC fragment thereof). The method is useful for RV detection and to assess
 CC graft transplantation risk. Screening of animals allows the elimination
 CC of donors with active replication of known viruses. Inactive proviruses
 CC can be detected and inactivated, allowing identification and elimination
 CC of potential human pathogens derived from swine in a manner not possible
 CC in the outbred human organ donor population and is important to the
 CC development of human xenotransplantation.
 XX Sequence 7333 BP; 1984 A; 1788 C; 1885 G; 1676 T; 0 other;
 XX
 alignment_scores:
 Quality: 2497.50 Length: 678
 Ratio: 4.328 Gaps: 13
 Percent Similarity: 85.103 Percent Identity: 71.829
 alignment_block:
 US-09-171-553B-6 x AAF77726 ..
 Align seg 1/1 to: AAF77726 from: 1 to: 7333
 4 ThrLeuSerArgArgHisLeuProThrArgGlyGluProLysArgLe 20
 |||||
 4757 ACCCTATCTCGTACTTTGGACCAACAGCGCTGTGAAGTCGA.... 4801
 20 uArg illeProLeuSerPheAlaSerIleAlaTrpPheLeuThr 36
 |||||
 4802 AGGATGCCCTTAAAGCTTCGCTCCATCGCGTGGTCTTACTGTCGA 4850
 37 IleTrpProGlnAlaSerSerLysArgLeuLeuAspSerSerAsnProH 53
 |||||
 4851 ATAACTCTCAAGTTAATGTAAGCGCTTGTGGACACCCGCAACTCCA 4900
 53 sArgProLeuSerLeuThrTrpLeuIleLeuAspProAspThrGlyValT 70
 |||||
 4901 TAAACCTATCTCGTACTGTGTTACTACTGCTCCGTTACAGGTATTA 4950
 70 hValAsnSerThrArgGlyValAlaProArgGlyThrTrpTrpProL 86
 |||||
 4951 ATATTAAAGACACTCAAGGAGGAGCTCCCTGGGAGCTGTGGCTTAA 5000
 87 LeuHisPheCysLeuArgLeuLeuAsnProAlaVal.....LysSerTh 101
 |||||
 5001 TTATATGTCTGCTCGATCAATATCCGTGTCATATGACCAAGCCAC 5050
 101 rProProAsnLeuValArgSerTyrGlyPheTyrCysCysProGly...T 117

|||||
 5051 ACCCCCGATGTACTCCGTGCTTACGGTTTACGTTGCCAGACCCC 5100
 117 hrcIuLysGluLysTyrCysGlySerGlyGluSerPheCysArgang 133
 |||||
 5101 CAATATATGAGAAATTTGTGGAAATCTCAGGATTTCTTTGGCAAGCAA 5150
 134 TrpSerCysValThrSerAsnAspGlyAspTrpLysTrpProIleSer 150
 |||||
 5151 TCGAGCTGCATTACTTCTAATGATGGGAATGTGAATGGCCACTCTCTCA 5200
 150 uGlnAspArgValLysPheSerPheValAsnSer..... 161
 |||||
 5201 GCAAGACAGAGTAAGTACTCTTTGTAAACATCTACCAAGTTATATAC 5250
 162GlyProGlyLysTyrLys.....MethLysLeuTyr 172
 |||||
 5251 AATTATATTATGGCATGGAGATGGAAGATTGGCAACAGCGGATCAA 5300
 173 LysAspLys.....SerCysSerProSerAspLeuAspTyr 184
 |||||
 5301 AAGATGTACGAATAAAGCAATAGCTGTCAATTCGTTAGACCTAGATTAA 5350
 184 rLeuLysIleSerPheThrGlu..ArgLysThrGlyLysTyrSerLysVa 200
 |||||
 5351 CTTAAATAATTAAGTTTCACTGMAAAAGAAACAGAAATATATCAAAAGT 5400
 200 LAspLysTrpTyrGluLeuGlyAsnSer..PheLeuLeuTyrGlyGly 216
 |||||
 5401 GGGTAATATGATATCTTGGGGAATATGTATC.....TATGAGGCTCT 5444
 217 Ala.....GlySerThrLeuThrIleArgLeuArgIleGluThrG 230
 |||||
 5445 GGGAGAAAGAAAGATCTCTTGTGACTATTCGGCTCAGATATAGAACTCA 5494
 5495 GATGGAACTCCGTTGCTATAGCAACAAATAGGGCTTGGCCGAAACAG 5544
 247 LyrProProAlaLeuGluProProHisAsnLeuProValProGlnLeuThr 263
 |||||
 5545 GACCTCAATCCAAAGAACAG..... 5564
 264 SerLeuArgProAspIleThrGlnProProSerAsnSerThrThrGlyLe 280
 |||||
 5565AGGCCATCTCTTAACCCCTCGATTACATAACACCTGTGATC 5608
 280 uIleProThrAsnThrProArgAsnSerProGlyValAlaProValLysThr 297
 |||||
 5609 AGTCCCACT.....GAGCCTAACATCACTACTATTAAACAG 5643
 297 LysGlnArgLeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSer 313
 |||||
 5644 GGGGAAACTTTTACCTCATCCAGGAGCTTTTCAAGCTCTTAACTCC 5693
 314 ThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyPr 330
 |||||
 5694 ACGACCTCAAGAGCTACCTCTTCTTGGCTTGTGCTTACCTTCGGGCC 5743
 330 oProTyrTyrGluGlyMetAlaLysGluArgLysPheAsnValThrLysG 347
 |||||
 5744 ACCTTACTATGAGGAGTGTAGAGGAGGAAATTTCAATGTACAAAGG 5793
 347 LysIleArgAsnGlnCysThrTrpLysSerArgAsnLysLeuThrLeuThr 363
 |||||
 5794 AACATTAAGACCAATGTACATGGGATCCCAAAATAAGCTTACCTTACT 5843
 364 GluValSerGlyLysGlyThrCysIleGlyLysAlaProProSerHisG 380
 |||||
 5844 GAGGTTCTGTGAAAGACACCTGATAGGATGTATCCCAATCCCAACCA 5893
 380 nHisLeuCysTyrSerThrValValTyrGluGlnAlaSerGluAsnGlnT 397
 |||||

```

5894 ACAACCTTTGTACACACAGTGAACCTTTAATGACACCTCTGAGAGTCAT 5943
397 YLLeuValProGlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThr 413
|||||
5944 ATCTGCTACCTGGTTATGACAGGTGGCGCATGTAATATCTGATTAAC 5993
414 ProCysValSerThrSerValPheAsnGlnSerLysAspPheCysValMet 430
|||||
5994 CCTTGTTTCCACCTTGGTTTTCACCAACCAACTAAAGACTTTTGGCTTAT 6043
430 tValGlnIleValProArgValIleTyrTrpHisProGlnGluValIleLeu 447
|||||
6044 GGTCCAAATTTGTCCTCCCGGGGTACTACTATATCCGAAAGACAGTCCCTG 6093
447 spGluTyrAspTyrArgTyrAsnArgProLysArgGlnProValSerLeu 463
|||||
6094 ATGAAATATGACTATAGATATTAATCGCCAAAAGAGAGCCCATATCCCTG 6143
464 ThrLeuAlaValMetLeuGlyLeuGlyThrAlaValAlaGlyValGly 480
|||||
6144 ACACGTAGCTGTAATGCTCGGATTGGAGTGGCTGCAGCGCTGGGAAACAG 6193
480 YThrAlaAlaLeuIleThrGlyProGlnGlnLeuGlnLysGlyLeuGly 497
|||||
6194 AACGGCTGCCCTAATCAGAGACCCCAACAGCTGAGAAAGAGACTTACATA 6243
497 ILeuHisAlaIleMetThrGlnAspLeuArgAlaLeuLysGlnSerVal 513
|||||
6244 ACCTACATCTGAATTTGTAAAGAGATCTCAAGCCCTTGAATAAAATCTCTC 6293
514 SerAsnLeuGlnGluSerLeuThrSerLeuSerGluValIleLeuGlnAs 530
|||||
6294 ACTTAACCTGGAGCAATCCCTAACCCTTATCTGAGAGTGGTTCTACAGAA 6343
530 nArgArgGlyLeuAspLeuPheLeuArgGlnGlyLeuGlyLeuGlyLeu 547
|||||
6344 CAGAAGGGGGTTAGATCTGTTATTTCTAAAGAGAGAGGTTATGCTAG 6393
547 IAlLeuLysGlnGluCysCysPheTyrValIleAspHisSerGlyAlaIleArg 563
|||||
6394 CCTTAAAGAGAGAAATGCTGCTTCTATGATGATCACTCAGAGCCATTCAGA 6443
564 AspSerMetAsnLysLeuArgLysLysLeuGlnArgArgArgGluArg 580
|||||
6444 GACTCATGAGCAACCTTAGAGAAAGTTAGAGAGCGCTCGAAGGGAAG 6493
580 GGLuAlaAspGlnGlyTTPheGlnGlyTTPheAsnArgSerProTirPM 597
|||||
6494 AGAGGCTGACAGGGGGTGGTTGAAGGATGGTTCACACAGGCTCTCTTGA 6543
597 eThrThrLeuLeuSerAlaLeuThrGlyProLeuValIleLeuLeu 613
|||||
6544 TGACCAACCTGCTTCTCTCTCTGACGGGGCCCTTAGTAGTCTCTCTCTG 6593
614 LeuLeuThrValGlyProCysLeuIleAsnArgPheValAlaPheVal 630
|||||
6594 TTACTTACAGTTGGGCTTCTCTTAAATTAATAGGTTTGTGCTTGTAG 6643
630 GGLuArgValSerAlaValGlnIleMetValLeuArgGlnGlnGlnGln 647
|||||
6644 AGAAGGAGTGTAGTCACTGACAGTCAATGAGTCAATGAGCAACAGTACCAAG 6693
647 LLeuLeuSerGlnGlyGluThrAspLeu 656
|||||
6694 GCCTTCGAGCCAGAGAAACTACCTC 6722

```

seq_name: /STDSD/gcgdata/geneseq/geneseqn/NA1997.DAT: AAT74883

seq_documentation_block:

ID AAT74883 standard: cDNA: 7393 BP.

XX AC AAT74883;

XX

```

DT 09-FEB-1998 (first entry)
XX Porcine retrovirus cDNA (defective).
DE Retrovirus; porcine; GAG protein; POL protein; ENV protein;
XX xenotransplantation; infectious; provirus; organ transplant; donor;
KW activated virus; PCR; ss.
XX Porcine retrovirus.
OS
FH Key Location/Qualifiers
FT CDS 598..2172
FT /tag= a
FT 598..2169
FT /tag= b
FT /note= "putative GAG protein"
FT CDS 2320..4737
FT /tag= c
FT /note= "putative POL coding region (partial) as
FT described in the specification"
FT mat_peptide 2320..3522
FT /tag= d
FT /note= "putative POL protein (partial)"
FT 3516..4328
FT /tag= e
FT /note= "putative POL protein (partial)"
FT 4332..4748
FT /tag= f
FT 4738..6725
FT /tag= g
FT /note= "putative ENV coding region (partial) as
FT described in the specification"
FT mat_peptide 4752..6722
FT /tag= h
FT /note= "ENV protein (partial)"
XX MO9721836-A1.
XX 19-JUN-1997.
XX 13-DEC-1996; 96WO-US19680.
XX 14-DEC-1995; 95US-0572645.
XX (GHEO ) GEN HOSPITAL CORP.
XX Fishman JA;
XX WPI: 1997-332804/30.
XX P-PSDB: AAM32091-W32095.
XX New nucleic acid from porcine retrovirus - used for detecting
XX viruses in transplant or other tissue and for assessing risk of
XX transmitting infection to graft recipient
XX Claim 16; Fig 2; 128bp; English.
XX
XX This cDNA sequence represents a defective purified swine retrovirus
XX found in PK-15 cells containing the putative coding regions for viral
XX GAG, POL, and ENV proteins. There are a few in frame stop codons and
XX apparent frame shifts in the given coding sequence which alter features
XX of the translation. This sequence and PCR fragments generated from the
XX sequence (see AAT74812-T74882) could be used to screen organs for the
XX presence of porcine retroviruses prior to xenotransplantation.
XX Transplantation can increase the likelihood of retroviral activation if
XX intact and infectious proviruses are present. The porcine retroviral
XX sequence can be used to generate probes to determine the level (e.g.
XX copy number) of intact (i.e. potentially replicating) porcine provirus
XX sequences in a strain of xenograft transplantation donors. It can be
XX used to detect mutations, genetic lesions or viral recombinants and to
XX determine the histological localisation of activated retroviruses. Using
XX Polymerase Chain Reaction DNA Quantitation (PQ) on blood mononuclear
XX cells, infectivity titration and susceptibility testing can be

```

CC performed. Ultimately animal donors without intact porcine retroviral
 CC sequences or a lower copy number of viral elements could be selected.
 XX
 SQ Sequence 7393 BP; 2004 A; 1801 C; 1896 G; 1692 T; 0 other;

alignment_scores:
 Quality: 2497.50 Length: 678
 Ratio: 4.328 Gaps: 13
 Percent Similarity: 85.103 Percent Identity: 71.829

alignment block:

US-09-171-553b-6 x AAT74883 ..

Align seg 1/1 to: AAT74883 from: 1 to: 7393

4 ThrLeuSerArgArgHisLeuProThrArgGlyGlyGluProLysArgLe 20
 |||||
 4757 ACCTTATCTGCTACTTTGACCCACACCAAGGCTGTGAAGTGA..... 4801
 20 uArg.IleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuThr 36
 |||||
 4802 .AGGAATCCCTTAAGCTTCGCCCTCCATCCGCTGCTTCTACTCTGTCA 4850
 37 IleThrProGlnAlaSerSerLysArgLeuIleAspSerSerAsnProH 53
 |||||
 4851 ATACCTCTCAAGTAAAGTAAACGCTGTGGACACGCCGCAACTCCCA 4900
 53 sArgProLeuSerLeuThrTrpLeuIleLeuAspProAspTrpGlyValT 70
 |||||
 4901 TAAACCTTATCTCTACCTGCTGTTACTTCTGACTCCGCTACAGGATTTA 4950
 70 hValAsnSerThrArgGlyValAlaProArgGlyThrTrpTrpProGlu 86
 |||||
 4951 AATATTAACAGACTCAAGGGGAGGCTCCCTGGGGACCTGTGGCTGAA 5000
 87 LeuHisPheCysLeuArgLeuLeuAsnProAlaVal..... LysSerTh 101
 |||||
 5001 TTATATGTCTGCCCTTCGATCATTAATCCCTGGTCTCAATGACCAAGCCAC 5050
 101 rProProAsnLeuValArgSerTrpGlyPheTrpCysCysProGly...T 117
 |||||
 5051 ACCCCCGAGTACTCCGCTTACGGGTTTACGTTTCCCAAGACCC 5100
 117 hGlyLysGluLysTrpCysGlyGlySerGlyLysPheCysArgArg 133
 |||||
 5101 CAAATATGAAGATATGTGGAATCTCCAGGATTTCTTTGCAAGCA 5150
 134 TrpSerCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLe 150
 |||||
 5151 TGGAGCTGCATACCTTCTAATATGAGGAATGGGAATGGCCAGTCTCA 5200
 150 uGlnAspArgValLysPheSerPheValAsnSer..... 161
 |||||
 5201 GGAAGACAGAGTAAGTACTCTTTTGTTRACATCTCTACCACTTAATC 5250
 162GlyProGlyLysTrpLys.....MetMetLysLeuTrp 172
 |||||
 5251 AATTTAATATGCGCATGGAGATGGAAAGATGGCAACAGGGGTACAA 5300
 173 LysAspLys.....SerCysSerProSerAspLeuAspTyr 184
 |||||
 5301 AAGAGATGACGAATAATAGCAATAATAGCTCATTCGTGACACATGATTA 5350
 184 rLeuLysIleSerPheThrGlu..ArgLysThrGlyLysTrpSerLysVa 200
 |||||
 5351 CTTAAATAATTAAGTCTGACTGAAAAAGAAACAGAAATAATTTCAAAAGT 5400
 200 LAspLysTrpTrpGlyLeuGlyAsnSer..PheLeuLeuTrpGlyGly 216
 |||||
 5401 GGGTAAATGATATATCTTGGGGAATAGGTAC.....TATGAGGGCTCT 5444
 217 Ala.....GlySerThrLeuThrIleArgLeuArgIleGluThrGlu 230

5445 GCGAAGAAAGAAAGATGCTGTCTGACTATATGCTCCAGAAATGAAACATCA 5494
 230 yThrGluProProValAlaMetGlyProAspLysValLeuIleGluGing 247
 |||||
 5495 GATGGAACTCGGTGTATAGGACAAATTAAGGTTGGCCGAACAG 5544
 247 lYProProAlaLeuGluProProHisAsnLeuProValProGluLeuThr 263
 |||||
 5545 GACCTTCATTCACAGAACAG..... 5564
 264 SerLeuArgProAspIleThrGlnProSerAsnSerThrThrGlyLe 280
 |||||
 5565AGGCCATCTCTTAACCCCTCTGATTACATACACCTCTGATC 5608
 280 uIleProThrAsnThrProArgAsnSerProGlyValProValLysTrpG 297
 |||||
 5609 AATCCCACT.....GAGCTTACATCTACTATTAACAG 5643
 297 lYGlArgLeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSer 313
 |||||
 5644 GGGCCAAACTTTTACCTCATCCAGSAGCTTTTCAAGCTCTTACTCTC 5693
 314 ThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyPr 330
 |||||
 5694 ACGACTCCAGAGGCTACTCTTCTTGTGGCTTGTAGCTTGCGGGCC 5743
 330 oProTrpTrpGluGlyMetAlaLysGluArgLysPheAsnValThrLysG 347
 |||||
 5744 ACCTTACTATGAGGAATGGCTAGAGAGGAAATTCATTAATGTGACAAAGG 5793
 347 lNHsArgAsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThr 363
 |||||
 5794 AACATAGAGACCAATGTATCATGGGATGCCAAATAATAGCTTACCCTTACT 5843
 364 GluValSerGlyLysGlyThrCysIleGlyLysAlaProProSerHisG 380
 |||||
 5844 GAGGTTCTTGGAAGGACCTGCATAGGAGTGGTCCCATCCACCA 5893
 380 nHisLeuCysTrpSerThrValValTrpGluGlnAlaSerGluAsnGlnT 397
 |||||
 5894 ACACCTTGTGACCACTGAAGGCTTAAATGCAACCTCTGAGAGTCAAT 5943
 397 yTrLeuValProGlyTrpAsnArgTrpTrpAlaCysAsnThrGlyLeuThr 413
 |||||
 5944 ATCTGTATCCGCTTATGACAGGCTGTGGGATGATTAATCTGAGTTAAC 5993
 414 ProCysValSerThrSerValPheAsnGlnSerLysAspPheCysValMe 430
 |||||
 5994 CTTGTGTCTCCACCTGTTGTTTCAACCAAACTAAAGACCTTTCGCTTAT 6043
 430 tValGlnIleValProArgValAlaTrpTrpHisProGluGluValValLeu 447
 |||||
 6044 GGTCCAAATGTGCCCGGGGTGATCTACTATCCGGAAGAACATCTCTG 6093
 447 sPgluTrpAspTrpArgTrpAsnArgProLysArgGluProValSerLeu 463
 |||||
 6094 ATGAATATGACTATAGATATATATGCGCAAAAGAGAGCCATATCTCCTG 6143
 464 ThrLeuAlaValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrG 480
 |||||
 6144 ACACACTGCTGAATGCTCGGATTTGGAGTGGCTGACAGGCTGGGAACAG 6193
 480 yThrAlaAlaLeuLeuThrGlyProGlnGlnLeuGluLysGlyLeuGlyG 497
 |||||
 6194 AACGGCTGCCCTTAATTCACAGGACCGCAACAGCTGGAAGAGGACTTAAT 6243
 497 lLeuHisAlaAlaMetThrGluAspLeuArgAlaLeuLysGlySerVal 513
 |||||
 6244 ACCTACATGCAATTTGTATAGGAGATGTCACAAACCTTGAAGAAATCTGTC 6293
 514 SerAsnLeuGluGluSerLeuThrSerLeuSerGluValValLeuGlnAs 530
 |||||

```

6294  AGAAGCCTGAGAGAAATCCCTAACCTCCTTATCATCGAAGTGCTTACAGAA  6343
530  nAArgatrgIyleuAspIleuLeuPheIleuArgIugIugIyleuCysIaa  547
6344  CAGAAAGGGGGTAGATCTTATTTCTTAAAAAGAGAGGGTTATGTGTAG  6393
547  IaIeulysIugIucYsCysPheTYIValAspHisSerGIalalIeArg  563
6394  CCTTAAAGAGGATCCTCCTCTATGTAGATCAGATCAGAGGACATAGA  6443
564  AspSerMetAsnIlysIleuArgIysIysIleuGluArgArgArgIuar  580
6444  GACTCCATAGCAAGAGCTTAGAGAAAGGTTAGAGAGCGCTGAAGGGAGAG  6493
580  gIuIuaIaAspGIugIyTrpPheGluIyTrpPheAsnArgSerProtrpm  597
6494  AAGAGGTGACCAAGGGGTGGTTGTAAGAGATGTTCAACAGGTCCTCTTGA  6543
597  eTrThrThrIleuLeuSerAlaIleuThrGlyProIeuValIleuLeuLeu  613
6544  TACACCACTCGCTTCTTGCTCTACGCGGGCCCTAGACAGTCCTGCTCG  6593
614  IeulLeuThrValGIyProCysIleuIleAsnArgPheValAlaPheVal  630
6594  TTACTACAGTGGCGCTTGCTTAATTAAGGTTTGGCTTTGTAG  6643
630  gGIuarGValSerAlaValGIleIleMetValIleuArgGIugIyTrgIug  647
6644  ACAAACAGTAGAGCGAGTCCAGATCATGTACTTAGGCAACAGTACCAAG  6693
647  IyleuLeuSerGIugIyGluThrAspIle  656
6694  GCCTTCTGAGCCAGAGGAATGACCTC  6722

seq_name: /SIBS2/ycgdata/geneseq/geneseqn/MA2001.DAT.AAC67023

```

CC the infectious agent, and then introducing these cells into the
CC transplant recipient.
XX
SQ Sequence 7873 BP; 2200 A; 1913 C; 1962 G; 1798 T; 0 other;

```

alignment_scores:
  Quality: 2474.50      Length: 704
  Ratio: 4.303          Gaps: 12
Percent Similarity: 81.676  Percent Identity: 67.330

```

```
alignment_block:
US-09-171-553B-6 x AAC67023
```

Align seg 1/1 to: AAC67023 from: 1 to: 7873

1	MeHisProThIleuSerArgrgHisIleuProFIhrArgIlyGlyIguLpr	17
5012	ATGCATCCACCGATTAAAGCGGCGCCACCTCCCGATTCGGGTGGAAAGCC	5061
17	oLyAsrGleuArgIleProIleuSerPheAlaSerIleAlaTrpPheLeuT	34
5062	GAAGAGACTGAAGATCCCTTAAGCTTCGCCATCCGCGGTGCTCTTA	5111
34	hIreuthrIleThrProGlnAlaSerSerIysArgIleuIleAspSerSer	50
5112	CTCTGTCATTAACCTCTCAAGTTAATGTAAACGCTTGTGTGACAGCCG	5161
51	AsnProHisArgProIleuSerLeuThrTrpIleuIleIleAspProAspTh	67
5162	AACTCCATAAACCCTTATCTCTCACTTGCTACTTACTGACCTCCGGTAC	5211
67	rgIyValThrValAsnSerThrArgIlyValAlaProArgIlyThrTrpT	84
5212	AGGTATTAAATTAACAGCACTCAAGGGAGGAGCTCCCTTGGGAGACTG	5261
84	rProGlnLeuHisPheCysIleuArgIleuIleAsnProAlaVal.....	98
5262	GGCTGATATTATGTCTGCTTCATGCATGCATGATCCGGTCTCAATGAC	5311
99	LySerThrProProAsnLeuValArgSerIyrgIyPheIyCysCysPr	115
5312	CAGCGCACACCCCGCATGTACTCGGTGTCAGGGCTTTACGTTTGCC	5361
115	ogLy.....ThGluLySGluIySerIyCysIyGlySerGlyIuSerPheC	131
5362	AGGACCCCCCAATTAATGAAGAAATATTGTGGAATCCTCAGATTTCTTT	5411
131	ysArGrArTrpSerCysValThrSerAsnAspGlyAspTrpIySerTrpPro	147
5412	GCAAGCAATGAGACTCGTAACTTAAATGATGGAAATTTGGAAATGGCCCA	5461
148	IleSerIleuGlnAspArgValIySPheSerPheValAsnSer.....	161
5462	GTCTCTCAGCAAGACAGAGTAAGTTACTTTTGTTAACAATCCTACCG	5511
162GlyProGlyIySerIyS.....MetMetL	170
5512	TTTAATACATTTAATTATGCGCAHGGAGATGGAAAGATTCGCAACAGC	5561
170	ysIleuTrpIyAspLys.....SerCysSerProSerAsp	181
5562	GGGTACAAAAGATGTACGAATAATMAACAATAGAGTCATTCGTTAAC	5611
182	LeuAspIyIleuLysIleSerPheThrIgu..ArgIySthrgIyIySyr	197
5612	CTAATATTACTTAATAATAGCTTCACTCGAAAAAGCAACAAGAAATAT	5661
198	SerIysValAspLysTrpIyrgIleuGlyAsnSer..PheLeuLeuTrpG	214
5662	TCAAAAGTGGTAAATGATGTGCTTGGGAAATAGGTAC.....TATG	5705
214	IyGlyGlyVala.....GlySerThrIleuThrIleArgIleuArgIle	227

5706 GAGCTCTGGAGAAAGATGTCGTGACATTCGCGCTCAGATA 5755
 228 GltThrGlyThrGluProProValAlaMetGlyProAspIysValLeuAl 244
 5756 GAAACCTCAGATGGAACTCCGGTCTGTATAGACCAATAAAGGTTGGC 5805
 244 agluginIglyProProAlaLeuGluProProHisAsnLeuProValProG 261
 5806 CGAACAGGAGCTCCAAATCCAGAACAG..... 5833
 261 lndleuThrSerLeuArgProAspIleThrGlnProProSerAsnSerThr 277
 5834AGGCCATCTCCTAACCCCTCGATTTACATACACACC 5869
 278 ThrGlyLeuIleProThrAsnThrProArgAsnSerProGlyValProVa 294
 5870 TCTGTGATCAGTCCCACT.....GAGCTTAACATCACTAT 5904
 294 llystrThrGlyInArgLeuPheSerLeuIleGlnGlyAlaPheGlnAlaI 311
 5905 TAAACAGAGGGCGGAAACTTTTAACTTCATCCAGGAGAGCTTTTCAGAGC 5954
 311 ldnSerThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSer 327
 5955 TTAACCTCCAGCTCCAGAGCTACCTCTCTTTGCTTGGCTTTGCTTAGCT 6004
 328 SerGlyProProTyrrTyrrGlyGlyMetAlaLysGluArgLysPheAsnVa 344
 6005 TCGGGCGCCACTTACTATGAGGGAATGGCTAGAGAGAGGAAATTCATGCT 6054
 344 lThrIysGluHisArgAsnGlnCysThrTrpGlySerArgAsnLysLeuT 361
 6055 GACAAAGGAGACATAGACACATGTACATGGGGATCCCAAAATTAACCTTA 6104
 361 hrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysAlaProPro 377
 6105 CCTTACGAGGTTTCTGGAAAAGACACGTGCATGAGGATGGTTCGCCCA 6154
 378 SerHisGlnHisLeuCysTyrrSerThrValValIlyrGlnAlaSerG 394
 6155 TCCCAACCAACACTTTGTAACCACTGAGACGCTTATATGAACTCTCA 6204
 394 uAsnGlnIlyrLeuValProGlyTyrrAsnArgTrpTrpAlaCysAsnThrG 411
 6205 GAGTACATCTCGTACTCGGTATGACAGTGTGGGCACTGAATATAGT 6254
 411 llyLeuThrProCysValSerThrSerValPheAsnGlnSerLysAspPhe 427
 6255 GATTAAACCTTGTGTTTCCACCTGTTTCAACCAACATAAGACTTT 6304
 428 CysValMetValGlnIleValProArgValIlyrTyrrHisProGlnIlyVa 444
 6305 TGCCTTATGTCCAAAATTGTCGCCGGGTGTACTACTATCCGAAAAAAC 6354
 444 lValLeuAspIuTyrrAspTyrrArgTyrrAsnArgProLysArgGluProV 461
 6355 AGTCTTGATGATATAGATAATAGATATATCGCCAAAAAGAGACCCA 6404
 461 alSerLeuThrLeuAlaValMetLeuGlyLeuGlyThrAlaValGlyVal 477
 6405 TATCCCTTACACTAGCTGTAATGCTCGAATGGGAATGGCTCGAGGCGTG 6454
 478 GlyThrGlyThrAlaAlaLeuAlaIleThrGlyProGlnGlyLeuGlyLysG 494
 6455 GGAACAGGAAGCGCTGCCCTAATACAGAGCCCAACACTGTGAGAAAG 6504
 494 lYleuGlyGlyLeuHisAlaAlaMetThrGlyuAspLeuArgAlaLeuLysG 511
 6505 ACTTAGTACCTACATCAATGTGAACGGAAAGATCTCAAGCCCTAGAAA 6554
 511 luserValSerAsnLeuGluGlnSerLeuThrSerLeuSerGluValVal 527

 6555 AATCTGTCACTAACCTGGAGGAATCCCTTACCTCTTATCTGAAGTGGTT 6604
 528 LeuGlnAsnArgArgGlyLeuAspLeuLeuPheLeuArgGluGlyLe 544
 6605 CTACAGAAACAGAGGGGGTACATCTGTATATTCTTAAAAAGAGAGATT 6654
 544 uCysAlaAlaLeuLysGluGluCysCysPheTyrrValAspHisSerGlyA 561
 6655 ATGTGTAGCCTTGAGAGAGGAATGTGTGTTTATGTGATCATTCAGGG 6704
 561 lalIleArgAspSerMetAsnLys..... 568
 6705 CCATCAGAGACTCCATGAACAGCTTAGAGAAAGACTGGAGAGCGCCCG 6754
 568 568
 6755 CGGCTCGAACAACCCAGACAGGTTCCTTGTTCATTAAGAACTGTC 6804
 569LeuArgLysLysLeuGluArgArg 577
 6805 GAAGTAACCGCTGAGCTAAAGCAGCTTAGAGAAAGTGGAGAAAGCGTC 6854
 577 rGArgGluArgGluAlaAspGlnGlyTrpPheGluGlyTrpPheAsnArg 593
 6855 GAAGGGAAAAAGAACTACTCAAGAGTGGTTGAGAGGATGGTTCAACAGG 6904
 594 SerProTrpMetThrThrLeuLeuSerAlaLeuThrGlyProLeuValVa 610
 6905 TCTCCCTTGGTGGCTACCTACTTCTGTGTTTACAGAGACCTTAATAGT 6954
 610 lLeuLeuLeuLeuLeuThrValGlyProCysLeuIleAsnArgPheValA 627
 6955 CTTCTCTCTGTACTACACAGTGGGCCATGATATATTAACAAGTTAATTG 7004
 627 lArPheValArgGluArgValSerAlaValGlnIleMetValLeuArgGln 643
 7005 CCTTCATTAGAGAACGAATTAAGTGCAGTCAGATCATGTGCTTAGACAA 7054
 644 GluTyrrGln 646
 7055 CAGTACCAA 7063
 seq_name: /SIDS2/9cdata/geneseq/geneseqn/NA1997.DAT:AAT74811
 seq_documentation_block:
 ID AAT74811 standard; cDNA; 8060 Bp.
 XX AAT74811;
 AC XX
 AC XX
 DT 11-FEB-1998 (first entry)
 XX
 DE Porcine retrovirus Tsukuba-1 cDNA.
 XX
 KW Retrovirus; porcine; GAG protein; POL protein; ENV protein;
 KW xenotransplantation; infectious; provirus; organ transplant;
 KW activated virus; Tsukuba-1; PCR; ss.
 OS Porcine retrovirus.
 XX
 FH Location/Qualifiers
 FH 86..2002
 FT /*tag= a
 FT /note= "putative ENV protein"
 FT 3112..4686
 FT /*tag= b
 FT /note= "putative GAG protein"
 FT 4871..8060
 FT /*tag= c
 FT /note= "putative POL protein (partial)"
 EN W09721836-A1.
 PD 19-JUN-1997.

XX 13-DEC-1996; 96MO-US19680.
 PF
 XX
 PR 14-DEC-1995; 95US-0572645.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 XX
 PI Fishman JA;
 XX
 DR WPI; 1997-332804/30.
 XX
 XX
 PT New nucleic acid from porcine retro:viruses - used for detecting
 PT viruses in transplant or other tissue and for assessing risk of
 PT transmitting infection to graft recipient
 XX
 PS Claim 1; Fig 1; 128pp; English.

This sequence represents the purified porcine retroviral cDNA sequence of Tsukuba-1 and contains the putative coding regions for viral proteins GAG, POL and ENV. This sequence and PCR fragments generated from the sequence (see AAT74812-T74882) could be used to screen organs for porcine retroviruses prior to xenotransplantation. Transplantation can increase the likelihood of retroviral activation if intact and infectious proviruses are present. The porcine retroviral sequence can be used to generate probes to determine the level (e.g. copy number) of intact (i.e. potentially replicating) porcine provirus sequences in a strain of xenograft transplantation donors. It can be used to detect mutations, genetic lesions or viral recombinants and to determine the histological localisation of activated retrovirus. Using Polymerase Chain Reaction DNA Quantitation (PDQ) on blood mononuclear cells, infectivity titration and susceptibility testing can be performed. Ultimately animal donors without intact porcine retroviral sequences or with a lower copy number of viral elements could be selected.

Sequence 8060 BP; 2233 A; 1959 C; 2012 G; 1856 T; 0 other;

```
alignment_scores:
  Quality: 2360.00      Length: 656
  Ratio: 4.252          Gaps: 11
Percent Similarity: 84.604  Percent Identity: 68.445
```

alignment_block:

Align seg 1/1 to: AAT74811 from: 1 to: 8060

[illegible][illegible]

```

408 sAsnThrGlyLeuThrProCysValSerThrSerValPheAsnGlnSerL 425
1258 TAATACGATGATTAACCCCTGTGTTCCACCTGGTTTAAACCAACATA 1307
425 ysaSpheCysValMetValGlnIleValProArgValIlyrTrHsPro 441
1308 AAGATTTTGCATTAATGTCACAAATGTGTCCTGAGTATTAATCTCC 1357
442 GlnGluValValLeuAspGluTrpAspTyrArgTyrAsnArgProLysAr 458
1358 GAAAAAGCAATCCTTGAATATGACTACAGAAATCATCGACAAAGAG 1407
458 gGluProValSerLeuThrLeuAlaValMetLeuGlyLeuGlyThrAlaV 475
1408 AGAACCATATCTCTGACACTTCTGTGATGCTCGAGACTTGAGTGCAG 1457
475 alGlyValGlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeu 491
1458 CAGGTAGAGAACAGACAGAGCTGCTGTCACGGGACACAGCAGCTA 1507
492 GlnGlyGlyLeuGlyGluLeuHisAlaIleMetThrGluAspLeuArgL 508
1508 GAAACAGGACTTAGTAACCTACATGCAATTGTAAACAGAAAGATCTCCAGC 1557
508 aLeuLysGluSerValSerAsnLeuGlnGluSerLeuThrSerLeuSerg 525
1558 CCTAGAAATCTGTCAGTACCTGAGAGATCCCTAACCTCTTATCTG 1607
525 lValValLeuGlnAsnArgArgGlyLeuAspLeuPheLeuArgGlu 541
1608 AATGATCTCTACAGATAGAGAGGCTTATTTATTTCTAAAGAA 1657
542 GlyGlyLeuCysAlaIleLeuLysGluGlyCysPheTyrValAspHi 558
1658 GAGAGATTATGTGTACCTTGAAGAGAGATCTGTTTATGTGATCA 1707
558 sSerGlyAlaIleArgAspSerMetAsnLysLeuArgLysLysLeuGluA 575
1708 TTCAGGGGCCATCAGAGACTCCATGAACTAGAGAAAGGTTGGAGA 1757
575 rGArgArgArgGluArgGluAlaAspGlnGlyTrpPheGlnGlyTrpPhe 591
1758 AGGTCGAGGAGAAAGAACTACTCAAGGCTGTTTCAAGGAGGCTTC 1807
592 AsnArgSerProTrpMetThrThrLeuLeuSerAlaLeuThrGlyProLe 608
1808 AACAGGTCTCTGTTGTTGCTACCTACTTCTGCTTACAGGACCTT 1857
608 uValValLeuLeuLeuLeuLeuThrValGlyProCysLeuIleAsnArgP 625
1858 AATAGTCCTCCTGCTGTTACTCAGATGTTGGCCATGTTATTATTAACAAGT 1907
625 heValAlaPheValArgGluArgValSerAlaValGlnIleMetValLeu 641
1908 TAATGCTCTCAATTAGAGAAAGATTAAGTGCAGTCCAGATCATGTTACTT 1957
642 ArgGlnGlnTyrGln 646
1958 AGACAAACAGTACCAA 1972

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF77725
seq_documentation_block:
ID AAF77725 standard; cDNA; 8060 BP.
XX
AC AAF77725;
XX
DT 23-MAY-2001 (first entry)
XX
DE Tsukuba-1 cDNA.
XX
KW Tsukuba-1; retrovirus; graft transplantation; xenotransplantation; ss.

```

```

XX OS Porcine retrovirus.
XX PN US6190861-B1.
XX PD 20-FEB-2001.
XX PF 13-DEC-1996; 9605-0766528.
XX PR 14-DEC-1995; 9505-0572645.
XX (GENE) GEN HOSPITAL CORP.
XX Fishman JA;
XX WPI; 2001-256211/26.
XX
XX PT Assessing risk of endogenous retroviruses in clinical practice and in
XX PT xenotransplantation, comprises using probe sequences derived from swine
XX PT or miniature swine retroviral genome.
XX
XX PS Claim 1; Fig 1; 127pp; English.
XX
XX CC The present invention relates to a method for screening a cell or tissue
XX CC for the presence or expression of a retrovirus (RV), comprising
XX CC contacting a target nucleic acid from the cell or tissue with a second
XX CC nucleic acid from the present invention (e.g. the present sequence or a
XX CC fragment thereof). The method is useful for RV detection and to assess
XX CC graft transplantation risk. Screening of animals allows the elimination
XX CC of donors with active replication of known viruses. Inactive proviruses
XX CC can be detected and inactivated, allowing identification and elimination
XX CC of potential human pathogens derived from swine in a manner not possible
XX CC in the outbred human organ donor population and is important to the
XX CC development of human xenotransplantation.
XX
XX S0 Sequence 8060 BP; 2233 A; 1959 C; 2012 G; 1856 T; 0 other;

alignment_scores:
Quality: 2360.00 Length: 656
Ratio: 4.252 Gaps: 11
Percent Similarity: 84.604 Percent Identity: 68.445

alignment_block:
US-09-171-553B-6 x AAF77725
Align seg 1/1 to: AAF77725 from: 1 to: 8060

1 MethisProThrLeuSerArgArgHisLeuProThrArgGlyGlyLupr 17
|||||
86 ATGCATCCACAGTTAAGCGGCGCCACCTCCGATTTGGGGGTGGAAGGCC 135
17 oLysArgLeuArgGlyLeuProLeuSerPheAlaSerIleAlaTrpPheLeu 34
|||||
136 GAAAGACTGAAATCCCTTAACCTTCGCTCCATGCGCTGTTGCTCTTA 185
34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSer 50
|||||
186 CTCTGTCAATTAACCTCTCAGACTAATGATGCGCATAGAGAGACGCTG 235
51 AsnProHisArgProLeuSerLeuThrTrpLeuIleLeuAspProAspTh 67
|||||
236 AACGCCATTAACCCCTTATCTCAGCTGTTATTAATGATGACGCGGAC 285
67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84
|||||
286 AGGTATTAAATATCAACAACTCAAGGGGAGGCTCTTTAGGAACCTGGT 335
84 rProGlnLeuHisPheCysLeuArgLeuLeuAsnProAlaValLysSer 100
|||||
336 GGCCTGATCTATACGTTGGCTCAGATCAGTTATTCCTAATCTGACCTCA 385
101 ThrProProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117

```

```

|||||..... 432
386 ...CCCCAGATATCTCCATGCTCAGGATTTATGTTGCCAGGACC 432
117 rGluYSerGlu...LysTYrCYSGlySergLYSerGluSerPheCYSAAGA 133
      :..... 433
433 ACCAATATATGAAACATTCGGGAATCCAGAGATTTCTTTGTAAAC 482
133 rGTrpSerCYsValThrSerAsnAspGlyAspTrpLysTrpProIleSer 149
      :..... 483
483 AATGGAAGCTGTACCTCTATGATGATATGGAATAGGCCAACCTCT 532
150 LeuGlnAspArgValLysPheSerPheValAsn.....SergLYProGlu 164
      :..... 582
533 CAGCAGGATAGGGTAAGTTTCTTATGTCACACCTATACCAAGCTCTGG 582
164 LysTYrLYSmetMetLysLeuTYrLysAspLysSer.....CysSerP 179
      :..... 632
583 ACAATTATATCTACCTGACCTGATTAAGAACGGAAGCCCAAGTCTCTC 632
179 rOSerAspLeuAspTYrLeuLysLleSerPheThr..GluArgLYSthrg 195
      :..... 682
633 CTTCAGACCTGATTAACCTAAAAATAAGTTTCACCTGAGAAAGAAACAA 682
195 LysTYrSerLYSValAspLysTrpTYrGluLeuGlyAsnSer..PheLe 211
      :..... 730
683 GAAATATCTCTAAATGGGTAATGTAATGCTTGGGGAATGTAAT.. 730
211 uLeuTYrGlyGlyGlyAla.....GlySerThrLeuThrIleArgL 225
      :..... 776
731 ...TATGAGGCTCGGGTAACACACAGGCTCCATCTAATCTATGCC 776
225 euArgIleGluThrGlyThrGluProProValAlaMetGlyProAspLys 241
      :..... 823
777 TCMAATATTAACCG...CTGAGCCTCCAAATGGCTATGAGCAACAAATCG 823
242 ValLeuAlaGluGlnGlyProProAlaLeuGluProProHisAsnLeuPr 258
      :..... 859
824 GTCTTGACGGCTCAAGAGACCCCAACCCAGAGACCA..... 859
258 oValProGluLeuThrSerLeuArgProAspIleThrGlnProProSerA 275
      :..... 872
860 .....GGACCATGCTCTA 872
275 snSerThrGlyLeuIleProThrAsnThrProArgAsnSerProGly 291
      :..... 916
873 ACATACACTCTGATCAGACCCCACTGATCTAGCAGCAGCACT..... 916
292 ValProValLysThrGlyGluArgLeuPheSerLeuIleGlnGlyAlaPh 308
      :..... 957
917 .....AATATGGGGGCAAACTTTTATGCTCATCCAGGAGCTTT 957
308 eGlnAlaIleAsnSerThrAspProAspAlaThrSerSerCysTrpLeu 325
      :..... 1007
958 TCAGCTCTTACTCCACGACTCCAGAGCTAAGCTCTTCTGTGTGGCTAT 1007
325 yLeuSerSerGlyProProTYrTYrGluGlyMetAlaLysGluArgLys 341
      :..... 1057
1008 GCTTAGCTTTGGGCCCTTACTATGAAAGGAGCTTAAAGAGGAAA 1057
342 PheAsnValThrLysGluHisArgAsnGlnCysThrTrpGlySerArgAs 358
      :..... 1107
1058 TTCATATGAGACAAAGAACATAGACACCAATGAGGATGCCAATAA 1107
358 nLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysA 375
      :..... 1157
1108 TAAAGCTTACCTTACTGAGGTTTCGAAAGGCACTGCATAGAAAGG 1157
375 laProProSerHisGlnHisLysCysTYrSerThrValValTYrGluGln 391
      :..... 1207
1158 TTCCCCCCTCCACCAACACCTTTGTATACACACCTGAAAGCCTTATATA 1207
392 AlaSerGluAsnGlnTYrLeuValProGlyTYrAsnArgTrpTrpAlaCy 408
      :.....

```

```

1208 ACCTGTAAGTCAATATCTGTACTGTTAGTATGACAGGTGGTGGCATG 1257
408 sAsnThrGlyLeuThrProCYsValSerThrSerValPheAsnGlnSerL 425
      :..... 1307
1258 TAATATGAGATTAACCCCTTGTGTTCACCTTGATTTTAAACCAACTA 1307
425 ySAspPheCYsValMetValGlnIleValProArgValTYrTYrHisPro 441
      :..... 1357
1308 AAGATTTTTCATATGATGTCAAATGTTTCCCGAGGTATTTACTATCC 1357
442 GluGluValAlaValLeuAspGluTYrAspTYrArgTYrAsnArgProLys 458
      :..... 1407
1358 GAATAACCAATCTCTGATGATATGACTACAGAAATCATGCAAAAAAG 1407
458 gGluProValSerLeuThrLeuAlaValMetLeuGlyLeuGlyThrAla 475
      :..... 1457
1408 AGAACCATATCTCTGACCTGCTGATGCTGAGCTCGGACTTGAGTGGCAG 1457
475 aLGIYValIGLYThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeu 491
      :..... 1507
1458 CAGGTGAGAGACAGCAAGACGCTGCCCTGGTCAGGGGACACAGCACTA 1507
492 GluLYSGlyLeuGlyGluLeuHisAlaAlaMetThrGluAspLeuArgAl 508
      :..... 1557
1508 GAACACAGACTTACTATACCTACATCGAATGTGACAGAAATCTCCAAAG 1557
508 aLeuLYSGlySerValSerAsnLeuGluGluSerLeuThrSerLeuSer 525
      :..... 1607
1558 CCTGAAATAATCTGTCTAGTAACCTGAGGAATCTTAACCTCTTACTG 1607
525 luValValLeuGlnAsnArgArgGlyLeuAspLeuPheLeuArgGlu 541
      :..... 1657
1608 AAGTAGTCCTACAGAAATGAAGAGGTTAGATTATTTATTTCTAAAGAA 1657
542 GlYLYLeuCYsAlaAlaLeuLYSGlyGluCYsCYsPheTYrValAspH 558
      :..... 1707
1658 GGAAGATTATGTGTAGCTTGAAGAGAGATGCTGTTTATATGTGATCA 1707
558 sSerGlyAlaIleArgAspSerMetAsnLYSLysLeuArgLYSLysLeuGlu 575
      :..... 1757
1708 TTCAGGGGCGCATCAGAGACTCCATGAACAACTTAGGAAGGCTTGAG 1757
575 rGArgArgArgGluArgGluAlaAspGlnGlyTrpPheGlyLysTrpPhe 591
      :..... 1807
1758 AGCCTCAAGAGGAAAGAACTACTCAAGGAGGTTTGAGGAGTGGTTC 1807
592 AsnArgSerProTrpMetThrThrLeuLeuSerAlaLeuThrGlyProLe 608
      :..... 1857
1808 AACGAGTCTCTGTGTGGCTACCTACTTTCGCTTAAACAGACCTT 1857
608 uValValLeuLeuLeuLeuLeuThrValGlyProCYsLeuIleAsnArgP 625
      :..... 1907
1858 AATGTCCTCTCTCTGTACTACAGTTGGGCATGATATTTAAACAAAG 1907
625 heValAlaPheValArgGluArgValSerAlaValGlnIleMetValLeu 641
      :..... 1957
1908 TAATGCTTCTTATGAGAAAGCAATAGTCACTCCAGATCATGTRACT 1957
642 ArgGlnGlnTYrGln 646
      :..... 1972
1958 AGACAAACAGTACCA 1972
seq_name: /sids2/gcgdata/geneseq/geneseqn/NA1997.DAT:AA74884
seq_documentation_block:
ID AA74884 standard; cDNA; 7892 BP.
XX
AC AA74884;
XX
DT 09-FEB-1998 (first entry)
XX
DE Miniature swine retrovirus cDNA.
XX

```

KW Retrovirus; porcine; GAG protein; POL protein; ENV protein;
 KW xenotransplantation; infectious; provirus; organ transplant; donor;
 XX activated virus; PCR; ss.
 OS Porcine retrovirus.
 XX
 FH Key Location/Qualifiers
 FT CDS 585..2159 /*tag= a
 FT mat_peptide 585..2156 /*tag= b
 FT /note= "putative GAG protein"
 FT CDS 2307..5744 /*tag= c
 FT mat_peptide 2307..5741 /*tag= d
 FT /note= "putative POL protein"
 FT CDS 5620..7536 /*tag= e
 FT mat_peptide 5620..7533 /*tag= f
 FT /note= "putative ENV protein"
 XX
 PN MO9721836-A1.
 PD 19-JUN-1997.
 PF 13-DEC-1996; 96WO-US19680.
 XX
 PR 14-DEC-1995; 95US-0572645.
 XX
 PA (GENO) GEN HOSPITAL CORP.
 XX
 PI Fishman JA;
 XX
 DR WPI; 1997-332804/30.
 DR P-PSDB; AAW32096-W32098.
 XX
 PT New nucleic acid from porcine retroviruses - used for detecting
 PT viruses in transplant or other tissue and for assessing risk of
 PT transmitting infection to graft recipient
 XX
 PS Claim 22; Fig 3; 128bp; English.
 XX
 CC This cDNA sequence represents a porcine retrovirus from miniature swine
 CC containing the putative coding regions for viral GAG, POL, and ENV
 CC proteins. This sequence and PCR fragments generated from it
 CC (see AAT74812-174882) can be used to screen organs for the presence of
 CC porcine retroviruses prior to xenotransplantation. Transplantation can
 CC increase the likelihood of retroviral activation if intact and
 CC infectious proviruses are present. The porcine retroviral sequence can be
 CC used to generate probes to determine the level (e.g. copy number) of
 CC intact (i.e. potentially replicating) porcine provirus sequences in a
 CC strain of xenograft transplantation donors. It can be used to detect
 CC mutations, genetic lesions or viral recombinants and also to determine
 CC the histological localization of activated retrovirus. Using Polymerase
 CC Chain Reaction DNA Quantitation (PQ) on blood mononuclear cells,
 CC infectivity titration and susceptibility testing can be performed.
 CC ultimately animal donors without intact porcine retroviral sequences or a
 CC lower copy number of viral elements could be selected.
 XX
 SQ Sequence 7892 BP; 2191 A; 1915 C; 1980 G; 1806 T; 0 other;
 alignment_scores:
 Quality: 2353.00 Length: 656
 Ratio: 4.232 Gaps: 12
 Percent Similarity: 84.756 Percent Identity: 68.598
 alignment_block:
 US-09-171-553B-6 x AAT74884 ..
 Align seg 1/1 to: AAT74884 from: 1 to: 7892

1 MethisProthleuSerArgArgHisleuProthArgglylgluPr 17
 |||||||||||||||||||||||||||||||||||||||||||
 5620 ATGCATCCCAACGTTAAACCGCGCCACCTCCGATTCGGGTGGAAGCC 5669
 17 olyArgleuArgllleProleuSerPheAlaSerlleAlatPrPheLeuT 34
 |||||||||||||||||||||||||||||||||||||||||||
 5670 GAAAGACGTGAAATCCCTTAAGCTTCCTCCATCGCGGTGTCCTTA 5719
 34 hrieuthrlleThrProGlnAlaSerSerlyArgleuIleAspSerSer 50
 |||||||||||||||||||||||||||||||||||||||||||
 5720 CTCTGTCAATACCTCTCAGACTAATGTATGCGCATAGGACAGCCCTG 5769
 51 AsnProHisArgProleuSerleuThrTrpLeuIleleAspProAspPh 67
 |||||||||||||||||||||||||||||||||||||||||||
 5770 AACTCCATTAACCCCTTATCTCTCAGCTGTATTAATTACGATCCGGCAC 5819
 67 rglYvalThrValAsnSerThrArgglyValAlaProArgglyThrTrpT 84
 |||||||||||||||||||||||||||||||||||||||||||
 5820 AGCTATTAAATATCAACACACTCAAGGGAGGCTCTCTTAGAACCTGTGT 5869
 84 rProGlnleuHisPheCysleuArgleuIleAsnProAlaValLysSer 100
 |||||||||||||||||||||||||||||||||||||||||||
 5870 GGCTGATCTATACGTTTGCCTCAGATCAATTCTCTAGTCTGACCTTA 5919
 101 ThrProAsnleuValArgSerTrpGlyPheTrpCysCysProglyTh 117
 |||||||||||||||||||||||||||||||||||||||||||
 5920 ..CCCCAGATATCCTCATGCTCAGGATTTTATGTTGGCCAGGACC 5966
 117 rglulysGlu...LysTrpCysglyLyserglyLysSerPheCysArg 133
 |||||||||||||||||||||||||||||||||||||||||||
 5967 ACCAAATTAATGGAACATTTGGGAAATCCAGAAATCTTTCTTTGTAAC 6016
 133 rTrpSerCysValThrSerAsnAspGlyAspTrpLysTrpProleser 149
 |||||||||||||||||||||||||||||||||||||||||||
 6017 AATGGAAGTGTATACCTCTATAGATGATTTGGAAATGGCAACCTGT 6066
 150 LeuGlnAspArgValLysPheSerPheValAsn.....SerGlyProG 164
 |||||||||||||||||||||||||||||||||||||||||||
 6067 CAGCAGATAGAGGTAGTATTCTTATATGCAACACTATACAGAGCTGTG 6116
 164 yLysTrpLysMetMetLysLeuTrpLysAspLysSer.....CysSerP 179
 |||||||||||||||||||||||||||||||||||||||||||
 6117 ACAATTAAATTACCTGACCTGATTTAGAACTGGAAGCTGCTCTTC 6166
 179 rSerAspLeuAspTrpLysLysIleSerPheThr..GluArgTrpThg 195
 |||||||||||||||||||||||||||||||||||||||||||
 6167 CTTCAGACCTAGATTACCTTAATAATAGTTTCTACAGAAAGAAACAA 6216
 195 yLysTrpSerLysValAspLysTrpTrpGlyLysGlyLysSer..PheLe 211
 |||||||||||||||||||||||||||||||||||||||||||
 6217 GAAATATCTCTTAATAATGCGTAAATGTAATGCTGTGGGGAATGGTATAT 6264
 211 uLeuTrpGlyLysGlyAla.....GlySerThrleuThrIleArgL 225
 |||||||||||||||||||||||||||||||||||||||||||
 6265TATGAGGCTCGCGTAAACCAACAGGCTCATTAATTAATTCGCC 6310
 225 euArgglyleuThrTrpGlyThrGlyLysProProValAlaMetLysProAspLys 241
 |||||||||||||||||||||||||||||||||||||||||||
 6311 TCATAATTAACCG...CTGGAACCTCCATAGGCTATAGGACCAATATAC 6357
 242 ValLeuAlaGluGlnGlyProProAlaLeuGluProHisAsnLeuPr 258
 |||||||||||||||||||||||||||||||||||||||||||
 6358 GTCTTGACGGGTCAAGAGCCCAACCAAGAGCA..... 6393
 258 oValProGlnLeuThrSerLeuArgProAspIleThrGlnProProSerA 275
 |||||||||||||||||||||||||||||||||||||||||||
 6394GAGCCATCCTCTTA 6406
 275 snSerThrTrpGlyLeuIleProthAsnThrProArgsnSerProgly 291
 |||||||||||||||||||||||||||||||||||||||||||
 6407 ACATTAACCTTGATCAGACCCCACTGAGTCT.....AACAGC..... 6444


```

US-09-171-553B-6 x AAF77727 ...
Align seg 1/1 to: AAF77727 from: 1 to: 8132

1 MethisProthLeuSerArgArgHisLeuProThrArgLysGlyLysIupr 17
|||||.....
5620 ATGCATCCACAGTTAAACGGCGCCACCTCCCGATCGGGGTGGAAAGCC 5669
17 OlyArgLeuArgLLeuProLeuSerPheAlaSerIleAlaIleThrPheLeu 34
|||||.....
5670 GAAAGACTGAAATCCCTTAACCTTCGCTCCATCGCTGGTCTCTTA 5719
34 hLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50
|||||.....
5720 CTGTGCAATACCTCTCAGACTATGCTATGCCATAGAGACAGGCTG 5769
51 AsnProHisArgProLeuSerLeuThrTrpLeuIleIleAspProAspTh 67
|||||.....
5770 AACTCCCATTAACCTTATCTCTCAGCTGGTTATCTAGCTCCGCGCAC 5819
67 rGlyValIleValAsnSerThrArgGlyValAlaIleProArgGlyThrTrp 84
|||||.....
5820 AGGTATTAAATTCACCAACACTCAAGGAGGCGCTCTTGAAGAACCTGGT 5869
84 rProGlnLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
|||||.....
5870 GGCTGATCTATACCTTGGCTCAGATCAGTTATCTCTAGCTGACCTCA 5919
101 ThrProAsnLeuValArgSerTrpGlyPheTrpCysCysProGlyTh 117
|||||.....
5920 ...CCCCAGATATCCCTCATGCTCAGGATTTATGTTGCCAGAGCC 5966
117 rGlnLysGlu...LysTrpCysGlyGlySerGlyGluSerPheCysArg 133
|||||.....
5967 ACCAAATATAGTGAACATTCGGGAATCCCGAGATTCTTGTGTAAAC 6016
133 rGTProSerCysValThrSerAsnAspGlyAspTrpLysTrpProLysSer 149
|||||.....
6017 AATGGAACCTGTAACTCTTAATGATGATATTTGMAATGCGCAACCTCT 6066
150 LeuGlnAspArgValLysPheSerPheValAsn.....SerGlyProG 164
|||||.....
6067 CAGCGAGTAGGTAAAGTTTCTTATGTCAACACTATACCACTCTGG 6116
164 yLysTrpIleMetLeuLysLeuTrpLysAspLysSer.....CysSer 179
|||||.....
6117 ACAATTTAATTAACCTGACCTGATTAAGMACTGAAGGCCCAAGGCTCTC 6166
179 rSerAspLeuAspTrpLeuLysIleSerPheThr..GluArgLysThrG 195
|||||.....
6167 CTTCAGACCTTAGATTACCTAAATAAATAGTTTCACTGACAAAGAAACAA 6216
195 IlyLysTrpSerLysValAspLysTrpTrpGluLeuGlyAsnSer..Phe 211
|||||.....
6217 GAAATATCTCTAAATAATGGTAAATGATGTTCTGGGGAATGATATAT.. 6264
211 uLeuTrpLysGlyGlyAla.....GlySerThrLeuThrIleArgL 225
|||||.....
6265 ...TATGGAGGCTCGGGTAAACAACGAGCTCATCTTCACTTACTTGTGCC 6310
225 euArgLLeuGluThrGlyThrGluProProValAlaMetGlyProAspLys 241
|||||.....
6311 TCAAAATATAACAG...CTGAGGCTTCCAATGGCTATATAGCAAAATAGG 6357
242 ValLeuAlaGluGlnGlyProProAlaLeuGluProProHisAsnLeuPr 258
|||||.....
6358 GTCTTGAGGCGTCAAAAGACCCCAACCAAGACCA..... 6393
258 oValProGlnLeuThrSerLeuArgProAspLLeThrGlnProProserA 275
|||||.....
6394 .....GGACCACTCTCTA 6406
275 snSerThrThrGlyLeuIleProThrAsnThrProArgAsnSerProGly 291

|||||.....
6407 ACATTAACCTCTGGATCAGACCCCACTGAGTCT.....AACAGC..... 6444
292 ValProValLysThrGlyArgLeuPheSerLeuIleGlnGlyAlaIle 308
|||||.....
6445 ...ACACATTAATGGGGCAAAACCTTTTACCTCATCTACAGGACCTTT 6491
308 eGlnAlaIleAsnSerThrAspProAspAlaThrSerSerCysTrpLeu 325
|||||.....
6492 TCAGGCTCTTAACCTCCAGACTCCAGAGGCTTACTCTTCTTGTGGCTAT 6541
325 yLysLeuSerSerGlyProProArgTrpGlyGlyMetAlaLysGluArgLys 341
|||||.....
6542 GCTTACTCTGGGGCCACCTTACTATGAAGATGCTGTAAGAGAGGANA 6591
342 PheAsnValThrLysGluHisArgAsnGlnCysThrTrpGlySerArgAs 358
|||||.....
6592 TTCAATGTGCAAAAGACATAGAGACCAATGCACATGGGGATCCCAAA 6641
358 nLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysA 375
|||||.....
6642 TAAGCTTACCCTTACTGAGGTTTCTGGAAGGACCTGCATAGAAAG 6691
375 lAbProSerHisGlnHisLeuCysTrpSerThrValValIlyrGluGln 391
|||||.....
6692 TTCCCCCATCCCAACCAACCTTTGTAAACCACTGAACCTTATATCA 6741
392 AlAserGluAsnGlnTrpLeuValProGlyTrpAsnArgTrpPrpLacy 408
|||||.....
6742 ACCTGTGAGAGTCAATATCTGTAACCTGTGTATGAGAGGTGGGGCATG 6791
408 sAsnThrGlyLeuThrProCysValSerThrSerValPheAsnGlnSerL 425
|||||.....
6792 TAATACTGAGTTAACCCCTTGTGTCTTCCACTTGTGTTTAACCAAC 6841
425 yAspPheCysValMetValGlnIleValProArgValIlyrTrpHisPro 441
|||||.....
6842 AAGATTTTTCATATAGTGCACAAATGTTCCCGAGATGATATCTATCC 6891
442 GluGluValValLeuAspGluTrpAspTrpArgTrpAsnArgProLysAr 458
|||||.....
6892 GAAAGCAATCTCTGATGATGATGACTACAGAAATCTGCAACAAAGAG 6941
458 gGluProValSerLeuThrLeuAlaIleMetLeuGlyLeuGlyThrAla 475
|||||.....
6942 AGAACCATATCTCTGACACTTCTGTGATGCTGCGACTTGGAGTGGCAG 6991
475 aLysGlyLeuGlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeu 491
|||||.....
6992 CAGGTGAGAACAGAACAGCTGCCCTGGTACGGGAGCACAGCAGTA 7041
492 GluLysGlyLeuGlyGluLeuHisAlaIleMetThrGluAspLeuArgL 508
|||||.....
7042 GAAACAGAGCTTGTAGTAACTTCAATGCAATTTGTAGACAGAGATCTCAACG 7091
508 aLeuLysGluSerValSerAsnLeuGluLysLeuThrLeuThrSerLeuSerG 525
|||||.....
7092 CCTTGAATAATCTGTCTGATACCTGAGAGATCCCTAACCTCTTATCTG 7141
525 IuValValLeuGlnAsnArgArgGlyLeuAspLeuLeuPheLeuArgGlu 541
|||||.....
7142 AAGTAGCTCTACAGAAAGAGAGGGTATGATTTATTTATTTCAAAAGAA 7191
542 GlyGlyLeuCysAlaAlaLeuLysGluGluCysCysPheThrValAspH 558
|||||.....
7192 GAGAGATTTATGTATGCTTGAAGAGAGATGCTGTTTATATGTGATCA 7241
558 sSerGlyAlaIleArgAspSerMetAsnLysLeuArgLysLysLeuGluA 575
|||||.....
7242 TTTCAGGGGCATCAGAGACTCCATGAACAGCTTACAGAAAGTGTGAG 7291
575 rArgArgArgGluArgGluAlaAspGlnGlyTrpPheGluGlyTrpPhe 591
|||||.....

```

```

7292 AGCGTGAAGGAGAAAGAACTACTCAAGGCTGTTTACGATGCTTC 7341
592 AsnArgSerProTrpMetThrThrLeuLeuSerAlaLeuThrGlyProLe 608
7342 AACAGGCTCTTGTGGTACCTACTTCTGCTTTACAGGACCCCT 7391
608 uValValLeuLeuLeuLeuLeuThrValGlyProCysLeuLeuAsnArgP 625
7392 AATAGCTCTCTCTCTGTTACTCAGAGTGGCCATGTATTATTACAGAGT 7441
625 hValValAlaLeuValArgGluArgValSerAlaValGlnIleMetValLeu 641
7442 TATTGGCTTCATTAGAGAACGATTAAGTCAGTCCAGTCTGCTACTT 7491
642 ArgGlnGlnTyrGln 646
7492 AGACAAACAGTACCAA 7506

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AA251087
seq_documentation_block:
ID AA251087 standard; DNA; 2000 BP.
XX
AC AA251087;
XX
DT 05-JUN-2000 (first entry)
XX
DE Contiguous DNA of porcine endogenous retrovirus-D (PERV-D) env region.
XX
KW Porcine endogenous retrovirus; PERV-D; virucide; prevention; vaccine;
XX pig; diagnosis; infection; xenotransplantation; antibody; env region; ds.
OS Sus scrofa.
XX
PN WO200011187-A1.
XX
PD 02-MAR-2000.
XX
PF 18-AUG-1999; 99WO-US19053.
XX
PK 18-AUG-1998; 98US-0097015.
XX
PA (BIOT-) BIO TRANSPLANT INC.
XX
PI Banerjee PT, Patience C, Andersson GK;
XX
DR WPI; 2000-224704/19.
XX
PT Porcine retroviral PERV-D polypeptides for diagnosing porcine
XX PT retroviral infections in humans after xenotransplantation -
XX PS Example 4; Fig 11; 119pp; English.
XX
CC The present DNA is the contiguous sequence of porcine endogenous
CC retrovirus-D (PERV-D) env region. It is compiled from the DNA fragments
CC obtained from the 5' end, extended proline rich region and 3' end of
CC PERV-D env region. It is isolated from the porcine kidney, PK15 cell line
CC (ATCC No. CCL-33). PERV-D DNA has 79% homology to a portion of PERV-C and
CC has virucidal activity. PERV-D sequence is useful for prevention or
CC diagnosis of infection of human tissues by porcine retroviruses after
CC xenotransplantation procedures. PERV-D polypeptide may be used to produce
CC specific antibodies, that can be administered as vaccines to create
CC passive immunity. The DNA can be used as a hybridisation probe or primer
CC for isolation purposes.
XX
SQ Sequence 2000 BP; 667 A; 468 C; 366 G; 499 T; 0 other:

```

```

alignment_scores:
Quality: 2136.00 Length: 654
Ratio: 3.963 Gaps: 14
Percent Similarity: 82.416 Percent Identity: 64.526

```

```

alignment_block:
US-09-171-553B-6 x AA251087
Align seg 1/1 to: AA251087 from: 1 to: 2000

```

```

1 MethIleProHleuSerArgArgHisLeuProThraArgIlyGlyLupr 17
...
43 ATACATCCACAGTTAAACCGCGCCACCTCTGATTCAGGAGTAAAGCC 92
17 OlyArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeu 34
93 AAAAGACTAAATATCCCTTAACTTACCTTCATCATTGCTTCCTTA 142
34 hLeuThrIleThrProGlnAlaSerSerLysArgLeuLeuAspSerSer 50
143 CTGTCAATAAACCTCTCAGACATATGATGATGACATAGAGACAGCTG 192
51 AsnProHleuArgProLeuSerLeuThrTrpLeuIleLeuAspProAsp 67
193 AACTCCATTAACCCCTTATCTGACCTGGTATTACTGATCTGACAC 242
67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84
243 AGGTATTATATATCCACAGCCGCTGAGGGAGGCTCTTTAGAAACCTG 292
84 rProGluLeuHisPheCysLeuArgLeuLeuAsnProAlaValLysSer 100
293 GGCCTGATCTATATGTCGCTCAGATCAGATCATTCCTAGTCTGACCT 342
101 ThrProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117
343 ACC...CCAGATATCCCTCGCTGCTTACGGATTTTATGTTTCCAC 389
117 rGluLysGlu...LysTyrCysGlyGlySerGlyLysPheCysArgA 133
390 ACCAATATATGGAAGAACATATGGAATCCCTAGAGATTCTTTACAAAC 439
133 rGTPSerCysValThrSerAsnAspGlyAspTrpLysTrpProIleSer 149
440 AATGAGAGCTGTATACCTTAATGATGAAATGAAATGCGCAACCTCT 489
150 LeuGlnAspArgValLysPheSerPheValAsnSerGlyProGlyLys 166
490 CTGAGAGATAGGGTAAAGCTTTCTTATGTCAC...CCCAATACCA 533
166 rLysMetLeuLysLeu...TyrLysAspLysSer...CysSerProSer 181
534 CCGAGACCTGGAAGACATACAGTACGATTCCTGCTTTCCCTGAG 583
181 sPLeuAspTyrLeuLysIleSerPheThrGlu...ArgLysThrGlyLys 197
584 ACCTAGATTATCTTAATAATAGTTCCAGCAAAAAAACAAGCAAAAT 633
197 rYSerLysValAspLysTrpTyrGluLeuGlyAsnSerPheLeuLys 213
634 ATCTTAATAATGATTAATAGTATGCTGGGAATAAT...ATATTATAC 680
214 rGlyGlyGly...AlaGlySerThrLeuThrIleArgLeuArgIleG 228
681 AGGTGGGACAGACCAACGAGCTCCATTATACATCCGACCTTAATAAT 729
228 luthrGlyThrGluProProValAlaMetGlyProAspLysValLeu 244
730 ...AGCCAGCTAGAGCTCCATGCTATAGGAGCCGATCGGCTTAAAG 777
245 GluGlnGlyProProAlaLeuGluProProHisAsnLeuProValPro 261
778 GGTCAAGAACCCCAACCCAGACCATCTCTGATATA..... 816
261 nLeuThrSerLeuArgProAspIleThrGlnProProSerAsnSerThr 278
817 .....ACTTCTAATTAGACCCACCTAG.....TCTACACAGCAGCA 853

```



```

278 hrGlyLeu1LeuProThrAsnThrProArgAsnSerProGlyValProVal 294
||
854 CT..... 855
295 LysThrGlyGluArgLeuPheSerLeuIleGlnGlyAlaPheGlnAlaI 311
||||||| ..... 311
856 AAAACGGGACAAACCTTTTATGCTCATCCAGGAGCTTTCAAGCTCC 905
311 eAsnSerThrAspProAspAlaThrSerSerCysThrLeuCysLeuSerS 328
||||||| ..... 328
906 TAACCTCAGAGCTCCAGAGGCTACCTCTTGCTGCTTGCCTTAACCTT 955
328 eGlyProProTyrTyrGluGlyMetAlaLysGluArgLysPheAsnVal 344
||||||| ..... 344
956 CGGGCCACCTTACTATAGAAATAAGGCTAAAGAAATAATTCATGTG 1005
345 ThrLysGluHisArgAsnGlnCysThrTrpGlySerArgAsnLysLeuTh 361
||||||| ..... 361
1006 ACMAAAACATAGAGACCAATGTATCATGGGATCCCAAAATTAAGCTTAC 1055
361 rLeuThrGluValSerGlyLysGlyThrCysIleGlyLysAlaProProS 378
||||||| ..... 378
1056 CCTTACTGAGGTTTCTGGAAAGACACCTGCATAAAGGTTCCCAT 1105
378 eHisGlnHisLeuCysTyrSerThrValValTyrGluGlnAlaSerGlu 394
||||||| ..... 394
1106 CCCACCAACACCTTACAAACCACTGAAGCCTTAAATCAAAACCTGAG 1155
395 AsnGlnTyrLeuValProGlyTyrAsnArgTrpTrpAlaCysAsnThrG 411
||||||| ..... 411
1156 AGTCATATCTGTGTAAGCTGTATGACAGAGTGCTGATGTAATATAC 1205
411 yLeuThrProCysValSerThrSerValPheAsnGlnSerLysAspPheC 428
||||||| ..... 428
1206 ATTAAACCCCTGTGTTCACCTGTGTTTCACCAAACTAAAGCTTTT 1255
428 yValMetValGlnIleValProArgValTyrTyrHisProGluGluVal 444
||||||| ..... 444
1256 ACATATATGTCMAATGTCCCGAGATATATTACTATCCCAAGAAACA 1305
445 ValLeuAspGluTyrAspTyrArgTyrAsnArgProLysArgGluProVa 461
||||||| ..... 461
1306 ATTCGTGATGATATGATTACAGCAACCATCGACAAAGAAACCCAT 1355
461 lSerLeuThrLeuAlaValMetLeuGlyLeuGlyTyrAlaValGlyValG 478
||||||| ..... 478
1356 ATCCCTGACACTCGCAGTAATCTCGAGCTCGAGATGATTAACAGGTGA 1405
478 lYThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeuGluLysGly 494
||||||| ..... 494
1406 GAACAGGAACCTGACCTTATGTTACAGGACCTCAGACGCTAGAAACAGA 1455
495 LeuGlyGluLeuHisAlaAlaMetThrGluAspLeuArgAlaLeuLysG 511
||||||| ..... 511
1456 CTTAGTACCTACATCAATGTATACAGAAATCTCCAAAGCCCTAAAGAA 1505
511 uSerValSerAsnLeuGluGlnSerLeuThrSerLeuSerGluValValL 528
||||||| ..... 528
1506 ATCTGTAGTAACTCGAAATAATCCCTAACCTCTTATCTGAAGTAGTTC 1555
528 eGlnAsnArgArgGlyLeuAspLeuLeuPheLeuArgGluGlyGlyLeu 544
||||||| ..... 544
1556 TACAGAAATAAAAGGTTAGATTATATTTTAAAGAAAGAAAGAAATTA 1605
545 CysAlaAlaLeuLysGluGlyCysCysPheTyrValAspHisSerGlyAl 561
||||||| ..... 561
1606 TGTGTAGCCTTAAAGAGAAATGCTTTTATGTATGATCATTCACAGGGCC 1655
561 aIleArgAspSerMetAsnLysLeuArgLysLysLeuGluArgArgArg 578
||||||| ..... 578
1656 CATCAGAGACTCCATGACACAGCTTAAGAAAGGTTGAGAAACGCTGCA 1705
578 rGluArgGluAlaAspGlnGlyTrpPheGluGlyTrpPheAsnArgSe 594
||||||| ..... 594
1706 GGGAAAGGAAACCTTACTCAAGATGTTTAAAGATGTTCAACAGGTC 1755
594 rProTrpMetThrThrLeuLeuSerAlaLeuThrGlyProLeuValValL 611
||||||| ..... 611
1756 TCCTTGTTGGCTTACCCTTACTTTTACTTTAACAGGACCTTATATGTC 1805
611 euleLeuLeuLeuThrValGlyProCysLeuIleAsnArgPheValAla 627
||||||| ..... 627
1806 TCCTCTGTTACTACAGATTGGGCCATGTATATTAAAGTAATATGCC 1855
628 PheValArgGluArgValSerAlaValGlnIleMetValLeuArgGlnG 644
||||||| ..... 644
1856 TTCATTGAAACGAAATAGTGAGTCAGATCATGTGTTAGTACACACA 1905
644 nTyrGln 646
|||||||
1906 GTACCAA 1912

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2002, 03:02:54 ; Search time 624.99 Seconds

(without alignments)
4054.870 Million cell updates/sec

Title: US-09-171-553b-9

Perfect score: 2956

Sequence: 1 tgccttttagggtaggaac.....aaaaaaaaaaaaaaaaaa 2956

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 segs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

N_Geneseq_1101:*

- 1: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2896.8	98.0	8196	AAV09699	Porcine retrovirus
2	2896.4	98.0	8209	AAV09700	Porcine retrovirus
3	2889.4	97.7	3482	AAV82749	Pig endogenous ret
4	1970.8	66.7	1974	AAV09703	Porcine retrovirus
5	1882.8	63.7	7333	AAV77726	Defective retrovir
6	1809.6	61.2	7393	AAV74883	Porcine retrovirus
7	1563.4	52.9	2462	AAV82748	Pig endogenous ret
8	1552.2	50.4	4918	AAV67022	PERV env protein c
9	1489.2	48.5	8132	AAV77727	Nucleotide sequenc
10	1433.8	45.6	7873	AAV67023	PERV env protein c
11	1347.8		7892	AAV74884	Miniature swine re

12	1231.8	41.7	8060	AAV74811	Porcine retrovirus
13	1231.8	41.7	8060	AAV77725	Tsukuba-1 cDNA. P
14	1009.8	34.2	1980	AAV67018	PERV-1-15 env prot
15	973.8	32.9	2000	AAZ51087	Continuous DNA of
16	719.4	24.3	7362	AAV67019	PERV env protein c
17	710.4	24.0	1493	AAZ51086	3'DNA of porcine c
18	622.8	21.1	4402	AAV67020	PERV env protein c
19	600.2	20.3	3320	AAV09698	Porcine retrovirus
20	455	15.4	6312	AAV09691	Plasmid pBdelPASA
21	451.2	15.3	8088	AAZ45540	Complete nucleotid
22	451.2	15.3	8535	AAQ73731	GALV SEAVO genome.
23	446	15.1	2129	AAA46343	DNA encoding the e
24	433.8	14.7	2055	AAV89076	GALV envelope glyco
25	430	14.5	2129	AAA46344	DNA encoding the e
26	419.4	14.2	1998	AAA89080	GALV envelope glyco
27	419.4	14.2	2004	AAV66507	DNA sequence of a
28	372.2	12.6	8655	AAV6750	Mus dunali endogeno
29	357.8	12.1	10970	AAQ73735	Retro virus vector
30	343	11.6	698	AAZ51073	5'DNA of porcine e
31	313.4	10.6	1941	AAV83504	Amphotrophic MLV r
32	309.2	10.5	1923	AAV83507	Amphotrophic MLV r
33	308	10.4	2484	AAQ05253	Sequence of clone
34	308	10.4	8440	AAV13265	FelV F6A provirus
35	307.6	10.4	1917	AAV89075	Amphotrophic glyco
36	307.6	10.4	1998	AAV83509	Amphotrophic MLV r
37	307.2	10.4	1602	AAV89139	Canine retrovirus
38	307	10.4	1959	AAV89074	Amphotrophic MLV r
39	307	10.4	6028	AAV90689	Plasmid pBdelPASF
40	306.8	10.4	1944	AAV83508	Amphotrophic MLV r
41	303.8	10.3	1965	AAV09276	V/H recombinant pr
42	302.2	10.2	1965	AAV83506	Amphotrophic MLV r
43	302	10.2	1950	AAV83505	Amphotrophic MLV r
44	301.8	10.2	1926	AAV89077	Murine leukemia vi
45	301.8	10.2	1929	AAV83503	Amphotrophic MLV r

ALIGNMENTS

AAV09699	1	AAV09699	standard; DNA; 8196 BP.
AAV09699		AAV09699	
AC	AAV09699		
XX			
DT	19-MAY-1998	(first entry)	
XX			
DE	Porcine retrovirus DNA encoding, GAG, POL and ENV.		
XX			
KW	Porcine retrovirus; POEV; POL protein; ENV protein; GAG protein;		
KM	vaccine; diagnosis; xenotransplantation; prophylactic; therapeutic; ds.		
XX			
OS	Porcine retrovirus.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	576..2126	
FT		/*tag= a	
FT		/product= GAG protein	
FT		/note= "viral core polypeptide"	
FT	CDS	2143..5733	
FT		/*tag= b	
FT		/product= POL protein	
FT		/note= "polymerase peptide sequence as given in	
FT		specification"	
FT	CDS	5606..7576	
FT		/*tag= c	
FT		/product= ENV protein	
FT		/note= "envelope protein"	
PN	W09740167-A1.		
XX			
PD	30-OCT-1997.		
XX			

PF 18-APR-1997: 97WO-GB01087.
 XX 10-FEB-1997: 97GB-0002668.
 PR 19-APR-1996: 96GB-0008164.
 XX (IMUT-) IMUTRAN LTD.
 PA (OONE-) O-ONE BIOTECH LTD.
 PI Galbraith DN, Haworth C, Lees GM, Smith KT;
 XX WPI, 1997-535851/49.
 DR
 XX Polynucleotide encoding porcine retrovirus expression product -
 PT useful to develop products for use in vaccines, diagnosis and
 PS xeno-transplantation
 XX
 PS Claim 4: Fig 2: 69pp: English.
 CC This DNA sequence encodes the porcine retrovirus (POEV) virion core
 CC polypeptide (GAG), polymerase (POL) and envelope (ENV) proteins. These
 CC proteins can be used to develop viral vaccines, antisense nucleic acids,
 CC ribozymes and other antiviral agents. They can also be used in
 CC xeno-transplantation technology and as diagnostic tools.
 XX
 SQ Sequence 8196 BP; 2165 A; 2061 C; 2147 G; 1820 T; 3 other:

Query Match 98.0%; Score 2896.8; DB 18; Length 8196;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 2937; Conservative 3; Mismatches 10; Indels 6; Gaps 3;

QY 1 tgccttttaggttaggaacacccctggacagtttgggtgcacccctatgattctct 60
 DB 5246 tgccttttaggttaggaacacccctggacagtttgggtgcacccctatgattctct 5305
 QY 61 acggggagaccccccgttgtagaattctcttctacatagtgctgattctctt 120
 DB 5306 acggggagaccccccgttgtagaattctcttctacatagtgctgattctctt 5365
 QY 121 cccagccttgcctctcagtcacagcgctcagtggttgtagaagcagcgctgagc 180
 DB 5366 cccagccttgcctctcagtcacagcgctcagtggttgtagaagcagcgctgagc 5425
 QY 181 agcccgagagagctactcagagagagagagacttgcaagtccacatcgcttccaa 240
 DB 5426 agcccgagagagctactcagagagagagagacttgcaagtccacatcgcttccaa 5482
 QY 241 gagattcagctatgttagagcgcacccgtgcaagaaacctcgagactcgttgaaagag 300
 DB 5483 gagattcagctatgttagagcgcacccgtgcaagaaacctcgagactcgttgaaagag 5542
 QY 301 ctatctcgtacttctgacacacacacagcgctgtagaagtcgaagaaatctccactg 360
 DB 5543 ctatctcgtacttctgacacacacacagcgctgtagaagtcgaagaaatctccactg 5602
 QY 361 tcaatgatccacagtaagctgagcgccactcccgactcgggttgtagagcgcaaaaga 420
 DB 5603 tcaatgatccacagtaagctgagcgccactcccgactcgggttgtagagcgcaaaaga 5662
 QY 421 ctggaatcccttaagactcgcctcactcagctggttcttactacttaaaactcc 480
 DB 5663 ctggaatcccttaagactcgcctcactcagctggttcttactacttaaaactcc 5722
 QY 481 caggcagtagtaaacgcttatagacagctcgaaaccccatagacacttatccctacc 540
 DB 5723 caggcagtagtaaacgcttatagacagctcgaaaccccatagacacttatccctacc 5782
 QY 541 tgggtgatttatgacctgtatagcggtgtgtactgtaataatgacctgaggtgtgtctcc 600
 DB 5783 tgggtgatttatgacctgtatagcggtgtgtactgtaataatgacctgaggtgtgtctcc 6042
 QY 601 agagacccctggtgctgactgacttctgctccgactgtaaaccccggtgtaaa 660
 DB 601 agagacccctggtgctgactgacttctgctccgactgtaaaccccggtgtaaa 660

DB 5843 agagacccctggtgctgactgacttctgctccgactgtaaaccccggtgtaaa 5902
 QY 661 agcacacccctccaaactgctgtagtattggttctctctgctccagagcagagaaa 720
 DB 5903 agcacacccctccaaactgctgtagtattggttctctctgctccagagcagagaaa 5962
 QY 721 gagaatactgttgggggttctgagaaatcctctctgtagagaaatgagctgcacactcc 780
 DB 5963 gagaatactgttgggggttctgagaaatcctctctgtagagaaatgagctgcacactcc 6022
 QY 781 aacgataagacttggaataatgacgctctctccagagcgggtaaaatctctctgctc 840
 DB 6023 aacgataagacttggaataatgacgctctctccagagcgggtaaaatctctctgctc 6082
 QY 841 aattccggcccgagcaatgacaaatgataaactataaaagataagagctgtcccca 900
 DB 6083 aattccggcccgagcaatgacaaatgataaactataaaagataagagctgtcccca 6142
 QY 901 tcaagacttagattatctaaagataaagttcactgaaagaaagaaacaggaataatccaa 960
 DB 6143 tcaagacttagattatctaaagataaagttcactgaaagaaagaaacaggaataatccaa 6200
 QY 961 aagtgataaattggtatgagctgaggaatgatttttaataatattgagcgagagcaagg 1020
 DB 6201 aagtgataaattggtatgagctgaggaatgattttttt-attatattgagcgagagcaagg 6259
 QY 1021 tccactttaaccattcgcttagatagatagacagggagacagaaacccctgtgcaattgag 1080
 DB 6260 tccactttaaccattcgcttagatagatagacagggagacagaaacccctgtgcaattgag 6319
 QY 1081 cccgataaagtattgctgagaaagggcccgcccgagggacacccgataacttgcg 1140
 DB 6320 cccgataaagtattgctgagaaagggcccgcccgagggacacccgataacttgcg 6379
 QY 1141 gtcgcccaatlaacctcgtcgtgagcctgacataaacaacagccgttagaacaatgataccact 1200
 DB 6380 gtcgcccaatlaacctcgtcgtgagcctgacataaacaacagccgttagaacaatgataccact 6439
 QY 1201 ggaattgattctcacaacacagccttagaaatcccccaggttctctgtagaagacaggaag 1260
 DB 6440 ggaattgattctcacaacacagccttagaaatcccccaggttctctgtagaagacaggaag 6499
 QY 1261 agactctcagctcatccagagagcttccaaagcacaactccacacagccagatgccc 1320
 DB 6500 agactctcagctcatccagagagcttccaaagcacaactccacacagccagatgccc 6559
 QY 1321 acttctctgttctgcttctatccctcagggcctcctatlatgagggagatggtctaaa 1380
 DB 6560 acttctctgttctgcttctatccctcagggcctcctatlatgagggagatggtctaaa 6619
 QY 1381 gaaagaataatcaatgtgacacaaagacatgaaatcaatgtatagatgggtcccgaaat 1440
 DB 6620 gaaagaataatcaatgtgacacaaagacatgaaatcaatgtatagatgggtcccgaaat 6679
 QY 1441 aagcttaacctcactgaagtttccgggaagggagatgcatagagaaagcctcccaatcc 1500
 DB 6680 aagcttaacctcactgaagtttccgggaagggagatgcatagagaaagcctcccaatcc 6739
 QY 1501 caccacaacacttctgataagactgtgttattagacagggcctcagaaatcagatattta 1560
 DB 6740 caccacaacacttctgataagactgtgttattagacagggcctcagaaatcagatattta 6799
 QY 1561 gtaactgtgtataaacaaggtgtgtgagcagaaatcagatggttaacccctgtgttccacc 1620
 DB 6800 gtaactgtgtataaacaaggtgtgtgagcagaaatcagatggttaacccctgtgttccacc 6859
 QY 1621 tcaagcttaacaaatccaaagatttctgtgcatcgggccaatcgttcccgagatgac 1680
 DB 6860 tcaagcttaacaaatccaaagatttctgtgcatcgggccaatcgttcccgagatgac 6919
 QY 1681 taccatctgaggaagtggttctgtatgataatgactatcgttataacagcaacaaagaa 1740
 DB 6920 taccatctgaggaagtggttctgtatgataatgactatcgttataacagcaacaaagaa 6979

QY 1741 gaaccggtatcccttaacctagctgttaatgctcggatlaaggagcgccgtttggttaga 1800
 |||||
 Db 6980 gaaccgtatcccttaacctagctgttaatgctcggatlaaggagcgccgtttggttaga 7039
 QY 1801 acaaggagcagctgctcgtatcacagagccacagcagctagagaagaagacttggtgacta 1860
 |||||
 Db 7040 acaaggagcagctgctcgtatcacagagccacagcagctagagaagaagacttggtgacta 7099
 QY 1861 catcgcccatcagacagaagatctccgagccttaaggagctgttaagcaaccatagaagag 1920
 |||||
 Db 7100 catcgcccatcagacagaagatctccgagccttaaggagctgttaagcaaccatagaagag 7159
 QY 1921 tccctgactctctgtctgtaagtggttctacagaaccgaggggattagatgactctgcttt 1980
 |||||
 Db 7160 tccctgactctctgtctgtaagtggttctacagaaccgaggggattagatgactctgcttt 7219
 QY 1981 ctaagagaagtggtgtatgtgcagccttaagaagaatgctgtctctatgtatgacac 2040
 |||||
 Db 7220 ctaagagaagtggtgtatgtgcagccttaagaagaatgctgtctctatgtatgacac 7279
 QY 2041 tcaagagccatcagacagatccatgaacaagcttagaaaaagttagagagcgctgaaag 2100
 |||||
 Db 7280 tcaagagccatcagacagatccatgaacaagcttagaaaaagttagagagcgctgaaag 7339
 QY 2101 gaaagagagctcagaccaggggtggttgaagagtggttcaacagctccttgatgacac 2160
 |||||
 Db 7340 gaaagagagctcagaccaggggtggttgaagagtggttcaacagctccttgatgacac 7399
 QY 2161 accctgtctctgtctgagagggccctctagctgtcctgttacttaacatgtggg 2220
 |||||
 Db 7400 accctgtctctgtctgagagggccctctagctgtcctgttacttaacatgtggg 7459
 QY 2221 cctgtcttaataagatgtgtgtcctgtttagagaagagtggtgcagtcagacac 2280
 |||||
 Db 7460 cctgtcttaataagatgtgtgtcctgtttagagaagagtggtgcagtcagacac 7519
 QY 2281 atgttacttaggaacaagtaaccaagggcctctgagccaaggagaaactgacactagcct 2340
 |||||
 Db 7520 atgttacttaggaacaagtaaccaagggcctctgagccaaggagaaactgacactagcct 7579
 QY 2341 tccaggttctaagaattagaacttlaacagagaagaagtggggaatgaaagatgaaaa 2400
 |||||
 Db 7580 tccaggttctaagaattagaacttlaacagagaagaagtggggaatgaaagatgaaaa 7639
 QY 2401 tgcacctaaccctccacagaaccaggaagtataaagaagctctaaatgccccgaatt 2460
 |||||
 Db 7640 tgcacctaaccctccacagaaccaggaagtataaagaagctctaaatgccccgaatt 7699
 QY 2461 acaagaccgtgctgctgccaagtaaatagtagaaggtcacaacttctatgttccagggc 2520
 |||||
 Db 7700 mcaagaccgtgctgctgccaagtaaatagtagaaggtcacaacttctatgttccagggc 7759
 QY 2521 ctgtctatccctgagtaagaattaacaggaatgagttgactaaactcgcttatctgatt 2580
 |||||
 Db 7760 ctgtctatccctgagtaagaattaacaggaatgagttgactaaactcgcttatctgatt 7819
 QY 2581 ctgttaaaactgactgacacatagaagaattgattacacatgacagcccttagtgacct 2640
 |||||
 Db 7820 ctgttaaaactgactgacacatagaagaattgattacacatgacagcccttagtgacct 7879
 QY 2641 tctcaactgcaatctgtcaacttgcccagagagccacagatgagacctccgagacta 2700
 |||||
 Db 7880 tctcaactgcaatctgtcaacttgcccagagagccacagatgagacctccgagacta 7939
 QY 2701 tttaaaatgagtgtcccgagcgcgggcctcgatattttaaatgattgtccatg 2760
 |||||
 Db 7940 tttaaaatgagtgtcccgagcgcgggcctcgatattttaaatgattgtccatg 7999
 QY 2761 gaagcgggcctcgtcgtatattttaaaatgattgtgttgacgacagggcttgctgtaa 2820
 |||||
 Db 8000 gaagcgggcctcgtcgtatattttaaaatgattgtgttgacgacagggcttgctgtaa 8059

QY 2821 ccccatlaaaagctgtcccgatccgacactcggggccgagctcctacccctgtgtg 2880
 |||||
 Db 8060 ccccatlaaaagctgtcccgatccgacactcggggccgagctcctacccctgtgtg 8119
 QY 2881 taagactgtgggccccagcgcttggaataaaatcctctgtgttttgcatacaaaaa 2940
 |||||
 Db 8120 taagactgtgggccccagcgcttggaataaaatcctctgtgttttgcatacaaaaa 8179
 QY 2941 aaaaaaaaaaaaaa 2956
 |||||
 Db 8180 aaaaaaaaaaaaaa 8195

RESULT 2

AAV09700 standard; DNA; 8209 BP.

AAV09700;

20-MAY-1998 (first entry)

Porcine retrovirus DNA encoding the LTR and GAG, POL and ENV proteins.

Porcine retrovirus; POEV; POL protein; ENV protein; GAG protein;

viral core polypeptide; polymerase protein; envelope protein;

vaccine; diagnosis; xenotransplantation; prophylactic; therapeutic; ds.

Porcine retrovirus.

Key

misc_feature

LTR

misc_feature

CDS

CDS

CDS

misc_feature

protein_bind

protein_bind

protein_bind

protein_bind

protein_bind

protein_bind

protein_bind

protein_bind

protein_bind

protein_bind

protein_bind

protein_bind

protein_bind

protein_bind

protein_bind

Query Match	98.0%;	Score 2896.4;	DB 18;	Length 8209;
Best Local Similarity	99.4%;	Pred. No. 0;		
Matches 2939;	Conservative	0;	Mismatches 11;	Indels 6;
				Gaps
QY 1	tgcttttaaggttaggaagaacacccctggaagaatttgctgtgaccccatgatatgctct	60		
Db 5260	tgcttttaaggttaggaagaacacccctggaagaatttgctgtgaccccatgatatgctct	53129		
QY 61	acgggggagaccccccgctgtgtagaatattgctctgtacatagtctatgtgtcgtctt	120		
Db 5320	acgggggagaccccccgctgtgtagaatattgctctgtacatagtctatgtgtcgtctt	53799		
QY 121	ccgagccttgttcttcaagctcgaagcgctcgaagtggtgtgagggacaacagcgctgaagc	180		
Db 5380	ccgagccttgttcttcaagctcgaagcgctcgaagtggtgtgagggacaacagcgctgaagc	54339		
QY 181	agctcgggagagcctactcaggaagagagacttgcgaattccacaatcgcttccaaigtg	240		
Db 5440	agctcgggagagcctactcaggaagagagacttgcgaattccacaatcgcttccaaigtg	54566		
QY 241	gagatcaagtctatgttaagcgccacccgctgtgcaggaacctctgaagactcgtgtgaaagagac	300		
Db 5497	gagatcaagtctatgttaagcgccacccgctgtgcaggaacctctgaagactcgtgtgaaagagac	55566		
QY 301	cttatctcgtacttttgcacacacaaacgagctgtgaagctgaagaactctccactgtga	360		
Db 5557	cttatctcgtacttttgcacacacaaacgagctgtgaagctgaagaactctccactgtga	56166		
QY 361	tcacatgcatcccaagcttaagctgtgcgcgaaccttcgcgactcggggtgtgaagcgcgaaaga	420		
Db 5617	tcacatgcatcccaagcttaagctgtgcgcgaaccttcgcgactcggggtgtgaagcgcgaaaga	56766		
QY 421	ctgagatacccttaagcttcgcgctccacatgcgctgtgttcccttaaccttaaatctccc	480		
Db 5677	ctgagatacccttaagcttcgcgctccacatgcgctgtgttcccttaaccttaaatctccc	57366		
QY 481	cagcgccagctagaacgctctatagaacagctgtgaaccccccatagaaccttaacctctacc	540		
Db 5737	cagcgccagctagaacgctctatagaacagctgtgaaccccccatagaaccttaacctctacc	57566		
QY 541	tgctgctatattgacccctgatacggggtgttctactgttaaatagactcgaagtgtgtccct	600		
Db 5797	tgctgctatattgacccctgatacggggtgttctactgttaaatagactcgaagtgtgtccct	58566		
QY 601	agaaggaacccctgtgctgtgaactgcattcttcgctccgattgatataaocccgctgttaa	660		
Db 5857	agaaggaacccctgtgctgtgaactgcattcttcgctccgattgatataaocccgctgttaa	59166		
QY 661	agcacacctcccaacctagatccgtaattatgaggtttctatgtgtgccagcgacagagaaga	720		
Db 5917	agcacacctcccaacctagatccgtaattatgaggtttctatgtgtgccagcgacagagaaga	59766		
QY 721	gagaatactgtgtggggttctgtgggaaatccctctgtaggagatvgagctgcacctccc	780		
Db 5977	gagaatactgtgtggggttctgtgggaaatccctctgtaggagatvgagctgcacctccc	60366		
QY 781	aagatgtgagactgtgaatattgcgcgactctctccagagccgggtaaatattcccttttc	840		
Db 6037	aagatgtgagactgtgaatattgcgcgactctctccagagccgggtaaatattcccttttc	60966		
QY 841	aattccgccccgggacaagtacaatattgataaactataaagatataagactgtgtcccca	900		
Db 6097	aattccgccccgggacaagtacaatattgataaactataaagatataagactgtgtcccca	61566		
QY 901	tcgaacttagattatctaagaatagtlttcactgtgaaaaggaacaggaataattcaa	960		
Db 6157	tcgaacttagattatctaagaatagtlttcactgtgaaaaggaacaggaataattcaa	62144		
QY 961	aagtgagtaaatgttagactctggggaatagttttttaataatatggcggggagacagg	10200		
Db 6215	aagtgagtaaatgttagactctggggaatagttttttatataatggcggggagacagg	6273		

QY 1021 tcaacttaaccattcgcttagatagatagacgagacgaacccctgtgcaatggga 1080
Db 6274 tcaacttaaccattcgcttagatagatagacgagacgaacccctgtgcaatggga 6333
QY 1081 ccgataaagtaactgctgtaacagagggcccccgccttgagacacccgataactgtgcg 1140
Db 6334 ccgataaagtaactgctgtaacagagggcccccgccttgagacacccgataactgtgcg 6393
QY 1141 gtgcccataatcaactcgctgacgctgacataacacgacgcttagacaactatcaact 1200
Db 6394 gtgcccataatcaactcgctgacgctgacataacacgacgcttagacaactatcaact 6453
QY 1201 ggaattgatctccataacacacgcttagaactccccaagtgtctcctgttaagacagacag 1260
Db 6454 ggaattgatctccataacacacgcttagaactccccaagtgtctcctgttaagacagacag 6513
QY 1261 agactcttcagctcatccacagggagcttcccaagccataactccacgcacccctgatgcc 1320
Db 6514 agactcttcagctcatccacagggagcttcccaagccataactccacgcacccctgatgcc 6573
QY 1321 aactcctctgtgtgcttcttctatactccagggccctctattatgagggatggctaa 1380
Db 6574 aactcctctgtgtgcttcttctatactccagggccctctattatgagggatggctaa 6633
QY 1381 gaaagaaatccaatgtgacccaagagacatagaatcaatgtacatggtgggtcccgcaat 1440
Db 6634 gaaagaaatccaatgtgacccaagagacatagaatcaatgtacatggtgggtcccgcaat 6693
QY 1441 aaggttaacctctcgtgaagtctccgggaagggagacatgtcctaaagaaaagctcccccattcc 1500
Db 6694 aaggttaacctctcgtgaagtctccgggaagggagacatgtcctaaagaaaagctcccccattcc 6753
QY 1501 caccacaacctgtgctatgactgtgtgttaagacagacccacagaatactcagattta 1560
Db 6754 caccacaacctgtgctatgactgtgtgttaagacagacccacagaatactcagattta 6813
QY 1561 gtaccgtgttataacagagtggtgacatgcaactggtgttaacccctgtgtttccacc 1620
Db 6814 gtaccgtgttataacagagtggtgacatgcaactggtgtgttaacccctgtgtttccacc 6873
QY 1621 tcaagtctcaaccaatccaagaattgtgtgtacatgtgtccaatgttcccccagagtac 1680
Db 6874 tcaagtctcaaccaatccaagaattgtgtgtacatgtgtccaatgttcccccagagtac 6933
QY 1681 taacatcttgaaggaagtgtctctgtgaatgaatgaatctcgtatataccgacccaataaaga 1740
Db 6934 taacatcttgaaggaagtgtctctgtgaatgaatgaatctcgtatataccgacccaataaaga 6993
QY 1741 gaaccggtatcccttaacctgaactgttaatgtctcgattagagagcggttgcgtaga 1800
Db 6994 gaaccggtatcccttaacctgaactgttaatgtctcgattagagagcggttgcgtaga 7053
QY 1801 aacggagacaagctgcctctgatacagagacacacagctagagaaagaaactgtgtgtaga 1860
Db 7054 aacggagacaagctgcctctgatacagagacacacagctagagaaagaaactgtgtgtaga 7113
QY 1861 catgtggtgcatacagagaatctccgagccttaaaagagatctgtttagcaaccttagaagag 1920
Db 7114 catgtggtgcatacagagaatctccgagccttaaaagagatctgtttagcaaccttagaagag 7173
QY 1921 tccctgaactcttctgtctgaagtgttctacagaacccggaaggatagatctgtctt 1980
Db 7174 tccctgaactcttctgtctgaagtgttctacagaacccggaaggatagatctgtctt 7233
QY 1981 ctaagagaagtggtgtatgtgacgaccttaaaagaaagtgtgtctctatgtatgacac 2040
Db 7234 ctaagagaagtggtgtatgtgacgaccttaaaagaaagtgtgtctctatgtatgacac 7293
QY 2041 tcaagagacatcagagactccatgaaacagcttagaataaagttagagagcgctcgaaag 2100
Db 7294 tcaagagacatcagagactccatgaaacagcttagaataaagttagagagcgctcgaaag 7353
QY 2101 gaaagagagctgacacaggggtgtttagaagatggttcaacaggtctcctgtgatgcc 2160

Db 7354 gaaagagagctgacacaggggtgtttagaagatggttcaacaggtctcctgtgatgcc 7413
QY 2161 accctgtcttctgtctgtacagggcccttagtagtctcgtctcgttacttaactgtgg 2220
Db 7414 accctgtcttctgtctgtacagggcccttagtagtctcgtctcgttacttaactgtgg 7473
QY 2221 cctgtcttaattaatggtttgttctgtttagagaaacaggtgtgacatccaagatc 2280
Db 7474 cctgtcttaattaatggtttgttctgtttagagaaacaggtgtgacatccaagatc 7533
QY 2281 atgttacttagagaaacagtaacaaagcctctcgtacgacaaaggaactactcagct 2340
Db 7534 atgttacttagagaaacagtaacaaagcctctcgtacgacaaaggaactactcagct 7593
QY 2341 tcccagttctaagatagaactaataacaagaacaaagtgtgggaatgaaagatgaaaa 2400
Db 7594 tcccagttctaagatagaactaataacaagaacaaagtgtgggaatgaaagatgaaaa 7653
QY 2401 tgaaccttaacctccccaacccaggaagttaataaagctcttaatgtccccgaatt 2460
Db 7654 tgaaccttaacctccccaacccaggaagttaataaagctcttaatgtccccgaatt 7713
QY 2461 acagacctgtgctgtgcagtaaatagtagaaggtcaccacttctatgttccagagc 2520
Db 7714 acagacctgtgctgtgcagtaaatagtagaaggtcaccacttctatgttccagagc 7773
QY 2521 ctgtatctcgtgcttaagttaagatacaggaatgaagtgtacttaactcgtatcgtat 2580
Db 7774 ctgtatctcgtgcttaagttaagatacaggaatgaagtgtacttaactcgtatcgtat 7833
QY 2581 ctgtataactgtactgtgcacatagaagaattgtattacatgtgacagccttagtacta 2640
Db 7834 ctgtataactgtactgtgcacatagaagaattgtattacatgtgacagccttagtacta 7893
QY 2641 tctcaactgtcaactgtgtcaactctgtccagagacccacagatgtgacacccctcagagcta 2700
Db 7894 tctcaactgtcaactgtgtcaactctgtccagagacccacagatgtgacacccctcagagcta 7953
QY 2701 ttlttaaatgatgtgttcaacagagcggtgctcgtctcgtatatttaaatgatgtgtcatg 2760
Db 7954 ttlttaaatgatgtgttcaacagagcggtgctcgtctcgtatatttaaatgatgtgtcatg 8013
QY 2761 gagcgcggtctcgtatatttaaatgatgtgtttagacacagccttgtgtgaa 2820
Db 8014 gagcgcggtctcgtatatttaaatgatgtgtttagacacagccttgtgtgaa 8073
QY 2821 ccccatlaaagctgtcccgatccgactcgggacgagctcctacccctgtgtgtg 2880
Db 8074 ccccatlaaagctgtcccgatccgactcgggacgagctcctacccctgtgtgtg 8133
QY 2881 taagactgtggcccccacagcggttggaataaaatcctctgtcgtttagcataaaaaa 2940
Db 8134 taagactgtggcccccacagcggttggaataaaatcctctgtcgtttagcataaaaaa 8193
QY 2941 aaaaaaaaaaaaaa 2956
Db 8194 aaaaaaaaaaaaaa 8209

RESULT 3
AAV82749
ID AAV82749 standard; DNA; 3482 BP.

XX AAV82749;

XX 25-FEB-1999 (first entry)

DE Pig endogenous retrovirus (PERV)-B envelope (env) gene region.

XX Pig endogenous retrovirus: PERV-A; envelope protein: PERV-B; subtype: KW probe: detection: retrovirus; human tissue; xenotransplant; KW primary porcine tissue; human cell line; porcine cell line; ss.

XX OS Pig endogenous retrovirus.
 XX FH Key Location/Qualifiers
 XX CDS 911..2884
 FT /+tag= a
 FT /product= envelope-protein
 XX PN W09853104-A2.
 XX PD 26-NOV-1998.
 XX PF 18-MAY-1998; 98MO-GB01428.
 XX PR 16-MAY-1997; 97GB-0010154.
 XX PA (MEDI-) MEDICAL RES COUNCIL.
 XX PI Stoye JP, Weiss RA;
 XX WPI: 1999-045324/04.
 DR P-PSDB: AAW85452.
 XX
 PT Newly isolated nucleic acid probe capable of hybridising to either
 PT the PERV-A or PERV-B env gene, useful in the detection of
 PT retroviruses, and their subtypes, in a sample of porcine/human
 PT tissue
 PS Claim 3: Page 21-23; 36pp; English.
 XX
 CC The present sequence encodes a Pig endogenous retrovirus (PERV)-B
 CC envelope protein. PERV exists in two different subtypes, PERV-A and
 CC PERV-B. The differences are reflected in sequence divergence in the
 CC envelope genes. Probes and primers can be derived from the envelope
 CC (env) genes of PERV-A and PERV-B. The probes and primers are used in
 CC a method to detect retroviruses in a sample of porcine/human tissue.
 CC particularly primary porcine tissue and human cell lines that have been
 CC cultivated in the presence of a porcine cell line, or human tissue from
 CC a patient with a xenotransplant. Subtype of PERV in a sample containing
 CC one of the PERV env genes can also be determined.
 CC
 SO Sequence 3482 BP; 927 A; 854 C; 867 G; 834 T; 0 other;

Query Match 97.7%; Score 2889.4; DB 20; Length 3482;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 2913; Conservative 0; Mismatches 21; Indels 1; Gaps 1;

QY 1 tgcctttaaggttagaaccaccctggacagtttgggtgagcccttaagttgctt 60
 DB 548 tgcctttaaggttagaaccaccctggacagtttgggtgagcccttaagttgctt 607
 QY 61 acgggggaaccccccgttggtagaatgtgtctgtataatagtgctgtgtctt 120
 DB 608 acgggggaaccccccgttggtagaatgtgtctgtataatagtgctgtgtctt 667
 QY 121 ccaagcttgttctctaggtcaagagcgtcagtggtgtgagcaacgagcgtggaac 180
 DB 668 ccaagcttgttctctaggtcaagagcgtcagtggtgtgagcaacgagcgtggaac 727
 QY 181 agctccggagagcctactcagagagagacttgaagtccacatcctcccaagt 240
 DB 728 agctccggagagcctactcagagagagacttgaagtccacatcctcccaagt 787
 QY 241 gagattcagtcatagttagacgcacacgctgcaggaacactcgagactcgttgaagagac 300
 DB 788 gagattcagtcatagttagacgcacacgctgcaggaacactcgagactcgttgaagagac 847
 QY 301 ctatctcgtacttttgaccacacaaagcgtgtgaaagtgcgaagaaatccacctgga 360
 DB 848 ctatctcgtacttttgaccacacaaagcgtgtgaaagtgcgaagaaatccacctgga 907
 QY 361 tccatgcatcccaagtcgtgcgcacacctccgacccgggtgtagagagccgaaaga 420

DB 908 tccatgcatcccaagtcgtgcgcacacctccgacccgtggtgtagagagccgaaaga 967
 QY 421 ctgagaatcccttaagcttgcctcccaatgcctgtgttcttactactcaataactcc 480
 DB 968 ctgagaatcccttaagcttgcctcccaatgcctgtgttcttactactcaataactcc 1027
 QY 481 cagccagtagtaaacgcttataagacagctcgaacccccaatagacttccctacc 540
 DB 1028 cagccagtagtaaacgcttataagacagctcgaacccccaatagacttccctacc 1087
 QY 541 tgcctgattatgtgacccctgatacgggtgtcaactgtaaatagactcgaagtggtctcc 600
 DB 1088 tgcctgattatgtgacccctgatacgggtgtcaactgtaaatagactcgaagtggtctcc 1147
 QY 601 agagcagactgtgtgcctgaacttgcttctgcctccgatttgatgaaccccggtttaa 660
 DB 1148 agagcagactgtgtgcctgaacttgcttctgcctccgatttgatgaaccccggtttaa 1207
 QY 661 agcacacctcccaactagtcctcgtagttatgtgtctatgtctgcacagacagagaa 720
 DB 1208 agcacacctcccaactagtcctcgtagttatgtgtctatgtctgcacagacagagaa 1267
 QY 721 gagaaatctgtgggggttctggggaatccttctgtagaagatgagctgtgtaacctcc 780
 DB 1268 gagaaatctgtgggggttctggggaatccttctgtagaagatgagctgtgtaacctcc 1327
 QY 781 aagctgagagactggaatctgcgactcctctccagagccgggttaaatctccttctc 840
 DB 1328 aagctgagagactggaatctgcgactcctctccagagccgggttaaatctccttctc 1387
 QY 841 aattccggccggcgaaglacaaaatgtatgaactataataagaatagactgtcccca 900
 DB 1388 aattccggccggcgaaglacaaaatgtatgaactataataagaatagactgtcccca 1447
 QY 901 tcaagcttagattactataagataagttcagttgaaaaaggaacccggaataattcaa 960
 DB 1448 tcaagcttagattactataagataagttcagttgaaaaaggaacccggaataattcaa 1507
 QY 961 aagtgataaatgtgataagctggtgggaatgttttataataatgtgctggggagcagg 1020
 DB 1508 aagtgataaatgtgataagctggtgggaatgttttataataatgtgctggggagcagg 1567
 QY 1021 tccacttaaccattgccttaagatagagaacggggagacacccctgtgtgcaatggga 1080
 DB 1568 tccacttaaccattgccttaagatagagaacggggagacacccctgtgtgcaatggga 1627
 QY 1081 cccgataaagtaactgtgcgaacaaaggcccccggccctggagccacgcataacttgcg 1140
 DB 1628 cccgataaagtaactgtgcgaacaaaggcccccggccctggagccacgcataacttgcg 1687
 QY 1141 gtcccccaattaacctctgtgcgtgtgacataaacacagccgcttaggaacagtaccact 1200
 DB 1688 gtcccccaattaacctctgtgcgtgtgacataaacacagccgcttaggaacagtaccact 1747
 QY 1201 ggtattgttcttaccacaaagccttagaaccctccaggtgttctgtttaaagacagagac 1260
 DB 1748 ggtattgttcttaccacaaagccttagaaccctccaggtgttctgtttaaagacagagac 1807
 QY 1261 agactctcagtcctatccagagagcttccaaagcactcaactccacagacccgtgagcc 1320
 DB 1808 agactctcagtcctatccagagagcttccaaagcactcaactccacagacccgtgagcc 1867
 QY 1321 acttctctgtgtgttctatccctcagggcctccttataatgagggagtgactaaa 1380
 DB 1868 acttctctgtgtgttctatccctcagggcctccttataatgagggagtgactaaa 1927
 QY 1381 gaaagaaatcctaatgtgacaaagagatagaaatcaatgtatcatgtgggttccgaaat 1440
 DB 1928 gaaagaaatcctaatgtgacaaagagatagaaatcaatgtatcatgtgggttccgaaat 1987
 QY 1441 aagcttaccctcactgaagtttccgggaaggagcaltgcataggaagaagctcccccctcc 1500

Db 1968 aagcttacctactgaagtttccgggaagggaacatgacatagaaagctccccctcc 2047
 QY 1501 caccacacacttgcctatagactggtttatagacgggctcagaagaatcacgtattta 1560
 Db 2048 caccacacacttgcctatagactggtttatagacgggctcagaagaatcacgtattta 2107
 QY 1551 gtacctgttatacaaggttggcgtacatgcaatctggttaacccccctgtttccacc 1620
 Db 2108 gtacctgttatacaaggttggcgtacatgcaatctggttaacccccctgtttccacc 2167
 QY 1621 tcaatgttcaacccaatccaagaatttgtgttcattgctccaaacgcccccgagttac 1680
 Db 2168 tcaatgttcaacccaatccaagaatttgtgttcattgctccaaacgcccccgagttac 2227
 QY 1681 tacatccctgaggaagtgtctctgataatatactatcgtataacccaagaaga 1740
 Db 2228 tacatccctgaggaagtgtctctgataatatactatcgtataacccaagaaga 2287
 QY 1741 gaaccggtatcccttaaccctagctgttaatgctcgatlaagagacgctgtgctaga 1800
 Db 2288 gaaccggtatcccttaaccctagctgttaatgctcgatlaagagacgctgtgctaga 2347
 QY 1801 acagggacagctgcccctgatacagagccacagacgtatagagaagaactgtgtgata 1860
 Db 2348 acagggacagctgcccctgatacagagccacagacgtatagagaagaactgtgtgata 2407
 QY 1861 catgcccgcatacagagaagatctccgagccttaagaagctgtgttaacacctaagaag 1920
 Db 2408 catgcccgcatacagagaagatctccgagccttaagaagctgtgttaacacctaagaag 2467
 QY 1921 tccctgactcttctgtcgaagtgtgtctacagaacccggagggatagatctgtgttt 1980
 Db 2468 tccctgactcttctgtcgaagtgtgtctacagaacccggagggatagatctgtgttt 2527
 QY 1981 ctaagagaaggttgggtatgtgcagccttaagaagaagatgtgtctctctatagatcac 2040
 Db 2528 ctaagagaaggttgggtatgtgcagccttaagaagaagatgtgtctctctatagatcac 2587
 QY 2041 tcaagagccatcagagatccatcagaaacagcttagaaaaaagttagagaagcgtcgaag 2100
 Db 2588 tcaagagccatcagagatccatcagaaacagcttagaaaaaagttagagaagcgtcgaag 2647
 QY 2101 gaaagaagaagctgacccaggggtgtgttgaaagatgtgtcaacagcttccctgagatgac 2160
 Db 2648 gaaagaagaagctgacccaggggtgtgttgaaagatgtgtcaacagcttccctgagatgac 2707
 QY 2161 accctgttctgtcgtcgaagggggccctagtgctcgtctcgttactactacagttggg 2220
 Db 2708 accctgttctgtcgtcgaagggggccctagtgctcgtctcgttactactacagttggg 2767
 QY 2221 ccttgcttaataatagatgtgtgtgctctgttagagaagaagtgagtgccagatc 2280
 Db 2768 ccttgcttaataatagatgtgtgtgctctgttagagaagaagtgagtgagtcagatc 2827
 QY 2281 atgtgtacttggaacagatcacagagcctctgaagccaagagaaactgaaccttagcct 2340
 Db 2828 atgtgtacttggaacagatcacagagcctctgaagccaagagaaactgaaccttagcct 2887
 QY 2341 tcccaagtctaaagttagaactatlaaacagacaagaagtgaggaaatgaaagatgaaaa 2400
 Db 2888 tcccaagtctaaagttagaactatlaaacagacaagaagtgaggaaatgaaagatgaaaa 2947
 QY 2401 tgaacaactaacctcccaagaccagaagttataaaaaagccttaaalggccccgaatt 2460
 Db 2948 tgaacaactaacctcccaagaccagaagttataaaaaagccttaaalggccccgaatt 3007
 QY 2461 acagagacctgtgctgccaagtaaatagtagaaggtcacacttctattgttccagggc 2520
 Db 3008 ccagagacctgtgctgccaagtaaatagtagaaggtcacacttctattgttccagggc 3067
 QY 2521 ctctatccctggcctaaagtagaataacagaagaatgtgtgactaatcgcttcttgatt 2580
 Db 3068 ctctatccctggcctaaagtagaataacagaagaatgtgtgactaatcgcttcttgatt 3127

QY 2581 ctgtaaaactgactggcaccatagagaagaattgattacacattgacagccctagtgaccta 2640
 Db 3128 ctgtaaaactgactggcaccatagagaagaattgattacacattgacagccctagtgaccta 3187
 QY 2641 tctcaactgcaalcctgtcaactctgcacagagagccacagatgcgagactccgagacta 2700
 Db 3188 tctcaactgcaalcctgtcaactctgcacagagagccacagatgcgagactccgagacta 3247
 QY 2701 ttcttaaaatgattgtgtccacagagagcggtcctcgatatttaaaatgattgttcacatg 2760
 Db 3248 ttcttaaaatgattgtgtccacagagagcggtcctcgatatttaaaatgattgttcacag 3307
 QY 2761 gagcgcggtc- tccatattttaaataattgtgtgtgtgagcgacagactgtgtgta 2819
 Db 3308 gagcgcggtc- tccatattttaaataattgtgtgtgtgagcgacagactgtgtgta 3367
 QY 2820 acccctaataaagctgtgccgattccgactcgggagcgagtcctacacccctgctgtg 2879
 Db 3368 acccctaataaagctgtgccgattccgactcgggagcgagtcctacacccctgctgtg 3427
 QY 2880 gtacgactgtgggccccagcgcttggaataaaatccctctgtgtgtgacatc 2934
 Db 3428 gtacgactgtgggccccagcgcttggaataaaatccctctgtgtgtgacatc 3482

RESULT 4

AAV09703

ID AAV09703 standard; DNA; 1974 BP.

XX AAV09703;

XX 20-MAY-1998 (first entry)

XX Porcine retrovirus Raji clone ENV DNA.

XX Porcine retrovirus: POEV; ENV protein; envelope protein; vaccine;

XX diagnosis; xenotransplantation; prophylactic; therapeutic; ds.

XX Porcine retrovirus.

XX Key Location/Qualifiers

XX CDS 1..1974

XX FT /*tag= a

XX FT /product= ENV protein

XX FT /note= "envelope protein"

XX W09740167-A1.

XX 30-OCT-1997.

XX 18-APR-1997; 97WO-G801087.

XX 10-FEB-1997; 97GB-0002668.

XX 19-APR-1996; 96GB-0008164.

XX (IMUT-) Q-ONE BIOTECH LTD.

XX (QONE-) Q-ONE BIOTECH LTD.

XX Galbraith DN, Haworth C, Lees GW, Smith KT;

XX WPI, 1997-535851/49.

XX P-PSDB; AAM39274.

This sequence encodes the porcine retrovirus envelope (ENV) protein isolated from the human cell line Raji. Such viral proteins can be used to develop viral vaccines, antisense nucleic acids, ribozymes and other

CC antiviral agents. They can also be used in xeno-transplantation
 CC technology and as diagnostic tools.

XX Sequence 1974 BP, 534 A; 489 C; 483 G; 468 T; 0 other;

Query Match 66.7%; Score 1970.8; DB 18; Length 1974;

Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1972; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 364 atgacatccacagttaagctgagccacacccacactccgactcgggtgagagagccgaaagactg 423
 DB 1 atgacatccacagttaagctgagccacacccacactccgactcgggtgagagagccgaaagactg 60
 QY 424 agaatcccttaagctgcgcctcatcgtcgtgttcttacttaacaataactcccg 483
 DB 61 agaatcccttaagctgcgcctcatcgtcgtgttcttacttaacaataactcccg 120
 QY 484 gccagtagtaaacgcttataagacagctcgaaaccccatagaccttatccctactcg 543
 DB 121 gccagtagtaaacgcttataagacagctcgaaaccccatagaccttatccctactcg 180
 QY 544 ctgattatgacctgtaacggtgltcaactgtlaaataagcaclcgagggtgtgtcctaga 603
 DB 181 ctgattatgacctgtaacggtgltcaactgtlaaataagcaclcgagggtgtgtcctaga 240
 QY 604 ggcacccgtggtgcctgtaacgcaatttctgctccgattgtaaccccggtgttaagc 663
 DB 241 ggcacccgtggtgcctgtaacgcaatttctgctccgattgtaaccccggtgttaagc 300
 QY 664 aacactcccaacactagctcgtatgattatggtttctatgctgcgcacaggaagaaagag 723
 DB 301 aacactcccaacactagctcgtatgattatggtttctatgctgcgcacaggaagaaagag 360
 QY 724 aaataactgtggtgttcttggtgaatccttctgtatagagatgagagctgtgcactccaac 783
 DB 361 aaataactgtggtgttcttggtgaatccttctgtatagagatgagagctgtgcactccaac 420
 QY 784 gatgagaactgtaaatgagcagatcctctccacagacgggtlaaatactctctgttcaat 843
 DB 421 gatgagaactgtaaatgagcagatcctctccacagacgggtlaaatactctctgttcaat 480
 QY 844 tcgagcccgagcaagtaacaaatgataaactatataagaataagagctctccccaata 903
 DB 481 tcgagcccgagcaagtaacaaatgataaactatataagaataagagctctccccaata 540
 QY 904 gactatagattatctaaagataaagtcttaactgaaagaaagaaacgggaaatatccaaag 963
 DB 541 gactatagattatctaaagataaagtcttaactgaaagaaagaaacgggaaatatccaaag 600
 QY 964 tggataaactgtaagactgaggaatagtttttaataatataagcggggagcagggtcc 1023
 DB 601 tggataaactgtaagactgaggaatagtttttaataatataagcggggagcagggtcc 660
 QY 1024 acttaacacatcgaccttagatagatagacggggagcaaacccctgtgccaatggagacc 1083
 DB 661 acttaacacatcgaccttagatagatagacggggagcaaacccctgtgccaatggagacc 720
 QY 1084 gataaactgactgctgtaacaaaggccccggcccttgagccaacagcctaacttcgggt 1143
 DB 721 gataaactgactgctgtaacaaaggccccggcccttgagccaacagcctaacttcgggt 780
 QY 1144 ccccaatlaacctgcgtcgagccgtgacataacacagccgcttagcaacagttaacactgta 1203
 DB 781 ccccaatlaacctgcgtcgagccgtgacataacacagccgcttagcaacagttaacactgta 840
 QY 1204 ttgattctaccacaacagccttagaataatcccaaggtgttctgtttaagacagagcagaga 1263
 DB 841 ttgattctaccacaacagccttagaataatcccaaggtgttctgtttaagacagagcagaga 900
 QY 1264 ctcttcaagctcatccaaaggagcttccaaagcctcaactccacggagccctgagtcact 1323
 DB 901 ctcttcaagctcatccaaaggagcttccaaagcctcaactccacggagccctgagtcact 960

QY 1324 tcttctgtgtgctgtgtctatcctcagggccctcttaattatgaaggatggtctaaagaa 1383
 DB 961 tcttctgtgtgctgtgtctatcctcagggccctcttaattatgaaggatggtctaaagaa 1020
 QY 1384 agaaatlaactgtgaccaaagagatagaaatcaatgtaactggtgtcccgaaatag 1443
 DB 1021 agaaatlaactgtgaccaaagagatagaaatcaatgtaactggtgtcccgaaatag 1080
 QY 1444 cttaaccttaactgaagttccgggaaggagacatcatagaaaggtcccccactccac 1503
 DB 1081 cttaaccttaactgaagttccgggaaggagacatcatagaaaggtcccccactccac 1140
 QY 1504 caacaccttgcatactgactgtgttataatgagcagcgccctcagaanaatcagattagta 1563
 DB 1141 caacaccttgcatactgactgtgttataatgagcagcgccctcagaanaatcagattagta 1200
 QY 1564 cctgtgtataacaggtgtgtggtgacatgcaatctggttaacccctgtgttccactca 1623
 DB 1201 cctgtgtataacaggtgtgtggtgacatgcaatctggttaacccctgtgttccactca 1260
 QY 1624 gtcttcaaccaaaccaaagattgtgtgtatgtgtccaaatcgtcccgaggtgtactac 1683
 DB 1261 gtcttcaaccaaaccaaagattgtgtgtatgtgtccaaatcgtcccgaggtgtactac 1320
 QY 1684 catcctgaaagaggtgtccttgatgataatgactaactgataaccgaccaaagagaga 1743
 DB 1321 catcctgaaagaggtgtccttgatgataatgactaactgataaccgaccaaagagaga 1380
 QY 1744 cccgtatcccttaacctagctgtatagtcgttaagtgaaagcggctgtggttgaagaa 1803
 DB 1381 cccgtatcccttaacctagctgtatagtcgttaagtgaaagcggctgtggttgaagaa 1440
 QY 1804 gggagacgtgcgcctgatcaacagacacagcagctagagaaagagactgtgtagtcact 1863
 DB 1441 gggagacgtgcgcctgatcaacagacacagcagctagagaaagagactgtgtagtcact 1500
 QY 1864 gcggccatgacagaagaatctccagaccttaagagactgtgttagcaacctagaagaatcc 1923
 DB 1501 gcggccatgacagaagaatctccagaccttaagagactgtgttagcaacctagaagaatcc 1560
 QY 1924 ctgactcttctgtcgaagtgtgttcttaacagacggaggggagattagatctgtgttcta 1983
 DB 1561 ctgactcttctgtcgaagtgtgttcttaacagacggaggggagattagatctgtgttcta 1620
 QY 1984 agagaaagtgtggttattgttcgaaccttaaaagaaatgtgtcttcatgtatagatca 2043
 DB 1621 agagaaagtgtggttattgttcgaaccttaaaagaaatgtgtcttcatgtatagatca 1680
 QY 2044 ggaagcacaatagagactccatgaacaaagcttagaanaaagttagaagagcgttcgaaggaa 2103
 DB 1681 ggaagcacaatagagactccatgaacaaagcttagaanaaagttagaagagcgttcgaaggaa 1740
 QY 2104 agagaaagtgtgacagaggtgtgttgaaggaatggttcaacaggtctccctgtgagccacc 2163
 DB 1741 agagaaagtgtgacagaggtgtgttgaaggaatggttcaacaggtctccctgtgagccacc 1800
 QY 2164 ctgacttctgcctcgaacgggcccactagactcgtctcctgttacttaacagtgtggcct 2223
 DB 1801 ctgacttctgcctcgaacgggcccactagactcgtctcctgttacttaacagtgtggcct 1860
 QY 2224 tgccttaataataggtctgtgtccttgtttaagaaacgggtgagtgaggttcaagtatgt 2283
 DB 1861 tgccttaataataggtctgtgtccttgtttaagaaacgggtgagtgaggttcaagtatgt 1920
 QY 2284 gtacttaagcaacagatcaacaaaggccttctgagcacaagagaaactgactctcag 2337
 DB 1921 gtacttaagcaacagatcaacaaaggccttctgagcacaagagaaactgactctcag 1974

RESULT 5
 AAF77726
 ID AAF77726 standard; DNA: 7333 BP.

XX AAF77726;
 AC 23-MAY-2001 (first entry)
 XX
 DT Defective retroviral genome isolated from PK-15 cell line.
 XX
 DE Retrovirus; graft transplantation; xenotransplantation; PK-15 cell line;
 XX ss.
 KW Unidentified.
 XX
 OS US6190861-B1.
 XX
 PN 20-FEB-2001.
 XX
 PD 13-DEC-1996; 96US-0766528.
 XX
 PF 14-DEC-1995; 95US-0572645.
 XX
 PR (GEHO) GEN HOSPITAL CORP.
 XX
 PA Fishman JA;
 XX
 PI WPI: 2001-256211/26.
 DR P-PSDB: AAB73282, AAB73283, AAB73284.
 XX
 XX Assessing risk of endogenous retroviruses in clinical practice and in
 PT xenotransplantation, complises using probe sequences derived from swine
 PT or miniature swine retroviral genome -
 XX
 PS Claim 1; Fig 2; 127pp; English.
 XX
 XX The present invention relates to a method for screening a cell or tissue
 CC for the presence or expression of a retrovirus (RV), comprising
 CC contacting a target nucleic acid from the cell or tissue with a second
 CC nucleic acid from the present invention (e.g. the present sequence or a
 CC fragment thereof). The method is useful for RV detection and to assess
 CC graft transplantation risk. Screening of animals allows the elimination
 CC of donors with active replication of known viruses. Inactive proviruses
 CC can be detected and inactivated, allowing identification and elimination
 CC of potential human pathogens derived from swine in a manner not possible
 CC in the outbred human organ donor population and is important to the
 CC development of human xenotransplantation.
 CC
 XX Sequence 7333 BP; 1984 A; 1788 C; 1885 G; 1676 T; 0 other;
 XX
 Query Match 63.7%; Score 1882.8; DB 22; Length 7333;
 Best Local Similarity 80.1%; Pred. No. 0;
 Matches 2409; Conservative 0; Mismatches 402; Indels 195; Gaps 9;

DB 4759 ctatctcgtacttttgaccacacacggtctgtgaagcgaag----- 4803
 QY 361 tcatgtatcccaagctgaagctgcccactcccgactcgggtgtgagagccgaaaga 420
 DB 4804 ----- 4803
 QY 421 ctgagaatcccttaagcttcgcctcatcgcctgtgtcttacttaacaataactcc 480
 DB 4804 -----gaatcccttaagcttcgcctcatcgcctgtgtcttacttaacaataactcc 4859
 QY 481 caggccaagtaagaagccttaagaagctcgaaccccatagaccttaactccatcc 540
 DB 4860 caagttaatgtgaacgccttctgtgcagcgcgaactcccatataacccttatctccac 4919
 QY 541 tggctattatctgaacctgaaggtgtcaactgttaataagaactcgaagtggtctcc 600
 DB 4920 tggtaacttaactgaactcgtacaggtatataataataagaagcactcaagggaggtccc 4979
 QY 601 agaggcaacctgtggtcgtgaactgcattctcgtcccgatgttaaccgcgtgta-- 658
 DB 4980 ttggggacctgtgtgctgtgaattatattgtcgtccttcgaatgaatccctgtctcaat 5039
 QY 659 -----aagaacacctcccaactagtcgtaagttatgtgtctgcgaag---c 711
 DB 5040 gaccaggaacacaccccgatgtaactcgtgtcgttaacggttttaagtttgcacagacc 5099
 QY 712 acggaagaagaagaatactacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 771
 DB 5100 ccaataatgaagaatattgtgtgaatccctcaggaatttcttgcagaagaatgtgtgt 5159
 QY 772 gtacctccaagatgtgagactgaagtgccatctctccaggaacgggttaaatc 831
 DB 5160 ataacttcaatgaatgtgaatttgaatgtgcagttccagcaagaagaatgaattac 5219
 QY 832 tctctgt-----caa 842
 DB 5220 tctttgttaacaatccctaccagttataatcaattatgtgcacatgtgtgtgtgtgt 5279
 QY 843 ttccgcccgggaagtaacaaatgtgaactataagaatgaagctgtcccatc 902
 DB 5280 gattgtcaacacggtgtacaaaagatg-taagaataagcaataagctgtcatcgt 5338
 QY 903 agacttagatttctaaagaataatcttcaataaagaagaagaagaagaatattcaaa 962
 DB 5339 agacttagatttctaaagaataatcttcaataaagaagaagaagaagaatattcaaa 5398
 QY 963 gtgataaattgtatgagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1019
 DB 5399 gt 5458
 QY 1020 gtccacttaacattcgtccttaagatagaagaggggaacgaacccctgtgtgcaatg 1079
 DB 5459 attgtgtcgaattatcgtccttaagatagaagatagaatgaatgaatcctgtgtcatag 5518
 QY 1080 acccgataaagtaactgt 1139
 DB 5519 accaataaaggt 5560
 QY 1140 ggtgtcccaataactcgt 1199
 DB 5561 -----acagaggtcattctccttaaccctcgtgtgtgtgtgtgtgtgtgtgtgt 5602
 QY 1200 tggattgtatcttaccacaagcctagaactcccgagtggtgtgtgtgtgtgtgtgtgtgt 1259
 DB 5603 tggattgtatcttaccacaagcctagaactcccgagtggtgtgtgtgtgtgtgtgtgtgt 5647
 QY 1260 ggaactctcaagctcaccagaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1319
 DB 5648 gaaactttttagccctacccagggaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5707
 QY 1320 caactctctgt 1379


```

FT      /*tag= 9
FT      /note= "putative ENV coding region (partial) as
FT      mat_peptide 4752..6722 described in the specification"
FT      /*tag= h
FT      /note= "ENV protein (partial)"
XX      W09721836-A1.
XX      19-JUN-1997.
XX      13-DEC-1996; 96WO-US19680.
XX      14-DEC-1995; 95US-0572645.
XX      (GEHO ) GEN HOSPITAL CORP.
XX      Fishman JA;
XX      WPI: 1997-332804/30.
XX      P-PSDB: AAM32091-W32095.
XX      New nucleic acid from porcine retroviruses - used for detecting
XX      PT viruses in transplant or other tissue and for assessing risk of
XX      PT transmitting infection to graft recipient
XX      Claim 16; Fig 2; 128pp; English.
XX      This CDNA sequence represents a defective purified swine retrovirus
XX      CC found in PK-15 cells containing the putative coding regions for viral
XX      CC GAG, POL and ENV proteins. There are a few in frame stop codons and
XX      CC apparent frame shifts in the given coding sequence which alter features
XX      CC of the translation. This sequence and PCR fragments generated from the
XX      CC sequence (see AAT7812-774882) could be used to screen organs for the
XX      CC presence of porcine retroviruses prior to xenotransplantation.
XX      CC Transplantation can increase the likelihood of retroviral activation if
XX      CC intact and infectious proviruses are present. The porcine retroviral
XX      CC sequence can be used to generate probes to determine the level (e.g.
XX      CC copy number) of intact (i.e. potentially replicating) porcine provirus
XX      CC sequences in a strain of xenograft transplantation donors. It can be
XX      CC used to detect mutations, genetic lesions or viral recombinants and to
XX      CC determine the histological localisation of activated retroviruses. Using
XX      CC Polymerase Chain Reaction DNA quantitation (PQ) on blood mononuclear
XX      CC cells, infectivity titration and susceptibility testing can be
XX      CC performed. Ultimately animal donors without intact porcine retroviral
XX      CC sequences or a lower copy number of viral elements could be selected.
XX      S0      Sequence 7393 BP; 2004 A; 1801 C; 1896 G; 1692 T; 0 other;
Query Match      61.2%; Score 1809.6; DB 18; Length 7393;
Best Local Similarity 78.5%; Pred. No. 0;
Matches 2407; Conservative 0; Mismatches 404; Indels 255; Gaps 10;
QY      1 tgcctttaaggcttagaacacaccccttgagacagtttgagctgacccctatgaattgctct 60
DB      4462 tgcctttaaggcttagaacacaccccttgagacagtttgagctgacccctatgaattgctct 4521
QY      61 acggggagacccccccctgtgtagaatgtctctgacatagagctgagtggtgtgttt 120
DB      4522 acggggagacccccccctgtgtagaatgtctctgacatagagctgagtggtgtgttt 4581
QY      121 cccaacctttctctctgagctcaagcgctcgagtggtgtaaggcaacgctggagac 180
DB      4582 cccaacctttctctctgagctcaagcgctcgagtggtgtaaggcaacgctggagac 4641
QY      131 agctccggagagctactacagagagagagagacttgcaattccacatcgctccaagtgtg 240
DB      4642 agctccggagagctactacagagagagagagacttgcaattccacatcgctccaagtgtg 4698
QY      241 gagatcagctatgtagacgacacgctgacgaaacactcgagactcggtggagagac 300
DB      4699 gagatcagctatgtagacgacacgctgacgaaacactcgagactcggtggagagac 4758

```

```

QY      301 ctatctcgtacttttgaccacacacacgagctgtgaaagtcgaaggaatctccaccttga 360
DB      4759 ctatctcgtacttttgaccacacacacgagctgtgaaagtcgaaggaatctccaccttga 4803
QY      361 tccatgcatccacglttaagctgagcgcaacccctcgagctgggtgtgagagcgcaaaaga 420
DB      4804 ----- 4803
QY      421 ctgagatcccttaagcttgctccatcgctcgtgttccctactctaacaataactcc 480
DB      4804 ----gaatcccttaagcttgctccatcgctcgtgttccctactctaacaataactcc 4859
QY      481 caggcagtagtaaacgctttagacagctgtaacccccatagacacttaactccctacc 540
DB      4860 caagttaaatggtaaagccttgtagcagcccggaactcccaataaaccttatctccacc 4919
QY      541 tggctgattatgacacctgtatagcaggtgtcactgttaaatagacacccgaggtgtctctc 600
DB      4920 tggctgattatgacacctgtatagcaggtgtcactgttaaatagacacccgaggtgtctctc 4979
QY      601 aagagcactgtgtgagctgaactgcatcttctgctccgattgattaaacccgctgta-- 658
DB      4980 ttgggagcactgtgtgagctgaactgcatcttctgctccgattgattaaacccgctgta-- 5039
QY      659 ----aagacacacccctccacactagtcgtagtattatggtgttcatgtgtccagg--c 711
DB      5040 gaccagcgcaacaccccgagtagctcgcgtgttaaggttttaagttgtccagacc 5099
QY      712 acagagaagaagaatactgtggtgtgtcgtgggaacccctctgtgagagatgtagctgc 771
DB      5100 ccaataatgaagaataatgtggaatccctcagagatttcttgcagaagaatgtagctgc 5159
QY      772 gtacactcaacagatgtagagagatgtagagatgtagagatgtagagatgtagagatgtag 831
DB      5160 ataacttcaatgtagagagatgtagagatgtagagatgtagagatgtagagatgtagagatgtag 5219
QY      832 tcccttgt-----caa 842
DB      5220 tctttgttaacactctacactcaagttataatcaatattatggtccatgtagatgtagaa 5279
QY      843 ttcggcgcggagaaatgaaatgtaagactataatgtaagtagtgcctccctacc 902
DB      5280 gattgcaacagaggggttaaaaaagatg--tacgaataagcaataaagcttcatctgct 5338
QY      903 agacttagattatcctaagaataagtttaactgtaaaaaaggaacacaggaataatcaaaa 962
DB      5339 agacttagattatcctaagaataagtttaactgtaaaaaaggaacacaggaataatcaaaa 5398
QY      963 gtgataaattgtagagctggtggaatgttttaataatataatggtggtgagc---agg 1019
DB      5399 gtgataaattgtagagctggtggaatgttttaataatataatggtggtgagc---agg 5458
QY      1020 gtccacttaacatcgtcgttagtagtagagagggagaaacccctgtgtagatgg 1079
DB      5459 atctgttctgacatcttcgtcctcagaataagaacatgagacagctcccggtgtgtagatgg 5518
QY      1080 acccgataaagtagctgctgaacaggggcccccggtcgtgtagacacacgcatatctgctc 1139
DB      5519 acccgataaagtagctgctgaacaggggcccccggtcgtgtagacacacgcatatctgctc 5560
QY      1140 ggtgcccacaaatlaactcgtcgtcgtcgtgaacacacacgctcgtgtagacacacg 1199
DB      5561 -----acagagcgcatctccttaacccctcgtgtagacacacacgctcgtgtagacacacg 5602
QY      1200 tggattgattctctacacacacgctcgtgtagacacacgctcgtgtagacacacgctcgtgtag 1259
DB      5603 tggattgattctctacacacacgctcgtgtagacacacgctcgtgtagacacacgctcgtgtag 5647
QY      1260 gaaactcttaagctcattccagagagcttccagacatcaactccacgacactgtagatgc 1319
DB      5648 gaaactcttaagctcattccagagagcttccagacatcaactccacgacactgtagatgc 5707

```


PA (MED-1) MEDICAL RES COUNCIL.

XX Stoye JP, Weiss RA;

XX WPI: 1999-045324/04.

DR P-PSDB: AAW85453.

XX Newly isolated nucleic acid probe capable of hybridizing to either
PT the PERV-A or PERV-B env gene - useful in the detection of
PT retroviruses, and their subtypes, in a sample of porcine/human
PT tissue

PS Claim 6; page 20-21; 36pp; English.

XX The present sequence encodes a Pig endogenous retrovirus (PERV)-A
CC envelope protein. PERV exists in two different subtypes, PERV-A and
CC PERV-B. The differences are reflected in sequence divergence in the
CC envelope genes. Probes and primers can be derived from the envelope
CC (env) genes of PERV-A and PERV-B. The probes and primers are used in
CC a method to detect retroviruses in a sample of porcine/human tissue,
CC particularly primary porcine tissue and human cell lines that have been
CC cultivated in the presence of a porcine cell line, or human tissue from
CC a patient with a xenotransplant. Subtype of PERV in a sample containing
CC one of the PERV env genes can also be determined.

XX Sequence 2462 BP; 715 A; 573 C; 575 G; 599 T; 0 other;

Query Match 52.9%; Score 1563.4; DB 20; Length 2462;

Best Local Similarity 79.5%; Pred. No. 0;

Matches 2000; Conservative 0; Mismatches 401; Indels 116; Gaps 8;

QY 151 tcgagtggtgaggaacagagcttggaagcagctccggagagcctactcagagagaag 210
DB 1 tcgagtggtgaggaacagagcttggaagcagctccggagagcctactcagagagaag 57
QY 211 acttgcagatccacatcgctcccaagtgtgagatcagctatgtttagagccagac 270
DB 58 acttgcagatccacatcgctcccaagtgtgagatcagctatgtttagagccagac 117
QY 271 caggaacactcagagactcgttggaagagacattatcgtactttgacacacaaag 330
DB 118 caggaacactcagagactcgttggaagagacattatcgtactttgacacacaaag 177
QY 331 ctgtgaaagtcgaagagatccacactgagatccatgacacagttgaagctggcgac 390
DB 178 ctgtgaaagtcgaagagatccacactgagatccatgacacagttgaagctggcgac 237
QY 391 ctccgcagctcggttgagagccgaaagagactggaatcccttaagcttcgcctcac 450
DB 238 ctccgcagctcggttgagagccgaaagagactggaatcccttaagcttcgcctcac 297
QY 451 gctcgttcttacttaacaataactccacagccagtaagtaaacgcttataagac 510
DB 298 gctcgttcttacttaacaataactccacagccagtaagtaaacgcttataagac 357
QY 511 tcgagaccatagactcttataccctactcgtgtgatttttagaccttgtagaggtgc 570
DB 358 tcgagaccatagactcttataccctactcgtgtgatttttagaccttgtagaggtgc 417
QY 571 acttaataagactcagagtggtctctcctagagagcactgtgtgcttgactatc 630
DB 418 acttaataagactcagagtggtctctcctagagagcactgtgtgcttgactatc 477
QY 631 tgcctcagattgaatacccgctgtta-----aaagcacactcccaactagtcgt 684
DB 478 tgcctcagattgaatacccgctgtta-----aaagcacactcccaactagtcgt 537
QY 685 agttatggtcttatcttctgcccag---cacagagaagaagatactgtgggtgc 741
DB 538 gcttaaggttttaacgtttgcccagagaccccaataatgaagatatttltggaacatc 597
QY 742 gggaatcctcttctgtagagatgagactcgtaaccctcaacagatgtagaactgaa 801

DB 598 caggaattcttcttgaacaaatgagcgtgacgaactcttaactgattggaattggaatg 657
QY 802 ccgactctctccagagccgggttaaaactctccttctgtaactc----- 846
DB 658 ccgactctctccagagcagagtaagtactcttcttgaacaaactcactacattat 717
QY 847 -----ggccggggaagtaacaatgatga 872
DB 718 caatttaattggtcattgagatgagatggaagatttggcaacagcggttcacaaagatg 776
QY 873 actataaagataagagagctctcccaactcagacttagattatctaaagataattcac 932
DB 777 acgaataatgaacaaataaagctgtacatctgttagacttagattacttaaaataattcac 836
QY 933 tgaanaagaaaacaggaataatctaaagtgtataatgtgtatgactgtgggaatgt 992
DB 837 tgaanaagaaaacaggaataatctaaagtgtataatgtgtatgactgtgggaatgt 896
QY 993 tttta-----ataatgctgggggagaggggtccacttaacactcgtccttagatga 1049
DB 897 gaactatgagagctctgagagagaagaagagatcttctgactatctgcctcagaatga 956
QY 1050 gacgggacagagacccctctgtgcaatgagaccgataaagtaactgctgagagggcc 1109
DB 957 aactcagatgagactcctcggtgtgataagaccataaaggttggcgaacagagacc 1016
QY 1110 cccgccttgagagccacgataacttgcggtgcccccaataactcgtcgtcgccgta 1169
DB 1017 tccaatccagagaca-----gagggcacc 1040
QY 1170 cataacacagccgcttagaagacagtaaccactgattgtatccacaaacagcctagaa 1229
DB 1041 tcttaaccccttgattacaataacacactcgtgactcagtcgccacttg----- 1088
QY 1230 ctcccaagtgcttctgttgaagacagagacagactcttcagctccatccagagactt 1289
DB 1089 ---gcctaacatcacatlaaaacagggcgagaaattttagccatccagggagactt 1145
QY 1290 ccaagcactcaactccacagccctgagtcactcttctgttggcttgccttcc 1349
DB 1146 tcaagctcttaactccacagactccagagctacactctctgttggcttgccttcc 1205
QY 1350 agggcctcttattagagggagtgctaaagaaagaaatcaattgtgaccaaagagca 1409
DB 1206 gggccacacttaactagagagagtgctagagaggaattcaattgtgaccaaagagca 1265
QY 1410 tagaatacagtgataggggtgtcccgaaataagcttaaccctcagtgagttccggaa 1469
DB 1266 tagaatacagtgataggggtgtcccgaaataagcttaaccctcagtgagttccggaa 1325
QY 1470 ggggagatgataaggaagagctcccccacccaacacacttctgtagtactgtgt 1529
DB 1326 aggcactgtgataaggaagagtggtcccccacccaacacacacttctgtagtactgtgt 1385
QY 1530 ttagagcagagcctcagaaatcgaatatttagtactgtgttataacaggttgggacatg 1589
DB 1386 cttaactgaaactctgagagatcgaatatttagtactgtgttataacaggttgggacatg 1445
QY 1590 caatctggtttaaaccctctgtgttccacactcagcttcaacaaatcgaatattgtg 1649
DB 1446 taactatgataatlaacccctgtgttccacactcagcttcaacaaatcgaatattgtg 1505
QY 1650 tgcattgtccaaatcgtcccgaggttactacacactccttggaaggttgccttgaatga 1709
DB 1506 cgttatggtccaaatgttcccgaggttactacacactccttggaaggttgccttgaatga 1565
QY 1710 atatgactatcgttataacccgaaagagaaagagacccgtatccctacactgtgtaac 1769
DB 1566 atagactatgataatcgtccgaaagagacccatccatccctgacactgtgtaac 1625
QY 1770 gctcgtatagagagcgtgtgtgtgtagaagagagacgtgccttgatcaacagagac 1829

Db 1626 gctcgattgagtgctgcagcgctgggaacaggaacgctgccttaacacagagac 1665
 QY 1830 acagagcttagaagaagacttgtagctacatgcgcacatgacagaagatctccgac 1889
 Db 1666 gcaacagctggagaagaagacttagtaacctacatgaattgtaacgaagatctccagc 1745
 QY 1890 cttaagagctgtgttaacacacttagaagatccctgtactcttctgtctgaagtgctt 1949
 Db 1746 cctagaaaaactctcagtaacctggaggaatccctacccctctatctgaagtgtct 1805
 QY 1950 acagaaccggagggatagatagatctgctgttcttaagaagaagtggttatctgacgct 2009
 Db 1806 acagacagaagaagggttagatctgtcttaattcttaagaagaagggttatgtgacgct 1865
 QY 2010 aaagaagaagaatgtgtctctctatgtatgacacagagacacacacacacacacac 2069
 Db 1866 aaagaagaagaatgtgtctctctatgtatgacacagagacacacacacacacacac 1925
 QY 2070 gcttagaaaaagttagaagaagcgctcgaaggagaagaaggctgacagagggtgttga 2129
 Db 1926 gcttagaaaaagttagaagaagcgctcgaaggagaagaaggctgacagagggtgttga 1985
 QY 2130 aggaatgttcaacagagctccttgatgacacacacacacacacacacacacacacac 2189
 Db 1986 aggaatgttcaacagagctccttgatgacacacacacacacacacacacacacacac 2045
 QY 2190 agtatgctcgtctctgttacttaacagttgggcttgccttaataatagttgttgcct 2249
 Db 2046 agtatgctcgtctctgttacttaacagttgggcttgccttaataatagttgttgcct 2105
 QY 2250 tgttagagaagaagatgtagtgcagtcacagatcagtgtagttagcaacaagacagac 2309
 Db 2106 tgttagagaagaagatgtagtgcagtcacagatcagtgtagttagcaacaagacagac 2165
 QY 2310 tcttagcgaaggagaacactgacacttagccttccagttctaagatgaagataataaca 2369
 Db 2166 tcttagcgaaggagaacactgacacttagccttccagttctaagatgaagataataaca 2225
 QY 2370 agacaagaagctgggaatggaatgaagaatgaagaatgaagaatgaagaatgaagaatga 2429
 Db 2226 agacaagaagctgggaatggaatgaagaatgaagaatgaagaatgaagaatgaagaatga 2285
 QY 2430 gttataaaaaagctctaaatgccccgaatatacagaacccctgctgctgctgctgctgct 2489
 Db 2286 gttataaaaaagctctaaatgccccgaatatacagaacccctgctgctgctgctgctgct 2345
 QY 2490 tagaaggtcgaacttccatctgttccagggcctgctatccctgctgctgctgctgctgct 2549
 Db 2346 tagaaggtcgaacttccatctgttccagggcctgctatccctgctgctgctgctgctgct 2405
 QY 2550 gaaatgagttgacttaactgctatctgctgctgctgctgctgctgctgctgctgctgct 2606
 Db 2406 gaaatgagttgacttaactgctatctgctgctgctgctgctgctgctgctgctgctgct 2462

RESULT 8

AAC67022

ID AAC67022 standard; DNA; 4918 BP.

XX AAC67022;

XX 27-MAR-2001 (first entry)

XX PERV env protein coding sequence SEQ ID NO: 22.

XX Xenotransplantation; infectious agent; vaccine; ds.

XX Porcine endogenous retrovirus.

XX MO200071726-A1.

XX 30-NOV-2000.

XX

PF 24-MAY-2000; 2000MO-US14296.
 XX
 PR 24-MAY-1999; 99US-0135631.
 XX
 PA (MAYO-) MAYO MEDICAL VENTURES.
 XX
 PI Federspiel MJ;
 XX
 DR WPI: 2001-032041/04.
 XX
 PT Inhibiting or preventing infectious agent transmission in mammalian
 PT transplant recipients, by introducing recombinant DNA comprising DNA
 PT encoding extracellular proteins of the agent into donor cells, such as
 PT swine cells -
 XX
 PS Claim 16; Page 109-111; 144pp; English.
 CC
 CC The present invention provides a method to prevent the transmission of
 CC infectious agents during xenotransplantation. This involves introducing
 CC to donor swine cells a recombinant DNA encoding a peptide fragment from
 CC the infectious agent, and then introducing these cells into the
 CC transplant recipient.
 CC
 XX Sequence 4918 BP; 1382 A; 1173 C; 1183 G; 1180 T; 0 other;

Query Match 52.5%; Score 1552.2; DB 22; Length 4918;
 Best local Similarity 73.2%; Pred. No. 0;
 Matches 2257; Conservative 0; Mismatches 603; Indels 223; Gaps 12;

QY 1 tgcctttagggtaggaacacccctgagacagttgggctgacccctatgaatgtct 60
 Db 1824 tgcctttagggtaggaacacccctgagacagttgggctgacccctatgaatgtct 1883
 QY 61 acgggggagaccccccccggtgtagaagaatgtcttctgtagatagtgctgtctt 120
 Db 1884 acgggggagaccccccccggtgtagaagaatgtcttctgtagatagtgctgtctt 1943
 QY 121 ccagaccccttctctcagcgaagcgctcgaatgggtgtagaagaagcgtagaagc 180
 Db 1944 ccagaccccttctctcagcgaagcgctcgaatgggtgtagaagaagcgtagaagc 2003
 QY 181 agctccggagagcctactcgaagaagaagacttgcaatccatcgtcccaagttg 240
 Db 2004 agctccggagagcctactcgaagaagaagacttgcaatccatcgtcccaagttg 2063
 QY 241 gagatcagctcagtttagagcgcacgcgtggaagaacactggaagactcgtggaaggac 300
 Db 2064 gagatcagctcagtttagagcgcacgcgtggaagaacactggaagactcgtggaaggac 2123
 QY 301 ctatctcgtacttttgacacacacacacagcgtgtggaagtcgaaggaatccacctgga 360
 Db 2124 ctatctcgtacttttgacacacacacacagcgtgtggaagtcgaaggaatccacctgga 2183
 QY 361 tccatgcatcccaagttgaagctgagccacccctcgcagctgggtgtagaagcgaagaaga 420
 Db 2184 tccatgcatcccaagttgaagctgagccacccctcgcagctgggtgtagaagcgaagaaga 2243
 QY 421 ctgagaatcccccctaagcttgcctcagctgagcgtgttcttactcttaacataacccc 480
 Db 2244 ctgagaatcccccctaagcttgcctcagctgagcgtgttcttactcttgaataactcct 2303
 QY 481 caggccagtagtaagcgtatagacagcgcgaaccccaataaccttaacccctaac 540
 Db 2304 caggccagtagtaagcgtatagacagcgcgaaccccaataaccttaacccctaac 2363
 QY 541 tgcctgattatgacccctgatacagcgtgtcactgttaaatgacactcgaagggtgtgtcct 600
 Db 2364 tgcctgattatgacccctgatacagcgtgtcactgttaaatgacactcgaagggtgtgtcct 2423
 QY 601 aagagcacctgtgagcctgaactcatttctgctcctgattgtaaccccgctgta-- 658
 Db 2424 ttgggagacctggtgctggaattatatagtctgctcctgcagtaagtaactcctgttcaat 2483

Db 4581 tgcgcacatgtgttaaccttaacaacatlaatttgcctccgagcgcggtccctc 4640
 QY 2736 gatatlaaaatgattgttccatgagcgcggtctcgcataattlaaattgatgct 2795
 Db 4641 gaagtttaattgactcgt-----ttgtatattttgaaatgatgct 4685
 QY 2796 tgtgacgacaggtctgtgttgtaaccccataaagctgtcccgatccgactcggtgc 2855
 Db 4686 tgltaagcgcggtctgtgttgtaaccccataaagctgtcccgatccgactcggtgc 4745
 QY 2856 cgcagctcctacccctcgtgtgtgtacgactgtggtcccgagcggtcggaataaaa 2915
 Db 4746 cgcagctcctacccctcgtgtgtgtacgactgtggtcccgagcggtcggaataaaa 4805
 QY 2916 tcctctgtgctgttgatcaaaa 2938
 Db 4806 tcctctgtgctgttgatcaaaa 4828

RESULT 9

AA77727 9
ID AAF77727 standard; DNA; 8132 BP.

AC AAF77727;

DT 23-MAY-2001 (first entry)

DE Nucleotide sequence of a retrovirus found in miniature swine.

KW Retrovirus; graft transplantation; xenotransplantation; miniature swine;

KW ss.

OS Unidentified.

PN US6190861-B1.

PD 20-FEB-2001.

PE 13-DEC-1996; 9605-0766528.

PR 14-DEC-1995; 9505-0572645.

PA (GEHO) GEN HOSPITAL CORP.

PI Fishman JA;

DR WPI; 2001-256211/26.

DR P-PSDB; AAB73285, AAB73286, AAB73287.

XX Assessing risk of endogenous retroviruses in clinical practice and in

PT xenotransplantation, comprises using probe sequences derived from swine

PT or miniature swine retroviral genome -

XX Claim 1; Fig 3; 127pp; English.

XX The present invention relates to a method for screening a cell or tissue

CC for the presence or expression of a retrovirus (RV), comprising

CC contacting a target nucleic acid from the cell or tissue with a second

CC nucleic acid from the present invention (e.g. the present sequence or a

CC fragment thereof). The method is useful for RV detection and to assess

CC graft transplantation risk. Screening of animals allows the elimination

CC of donors with active replication of known viruses. Inactive proviruses

CC can be detected and inactivated, allowing identification and elimination

CC of potential human pathogens derived from swine in a manner not possible

CC in the outbred human organ donor population and is important to the

CC development of human xenotransplantation.

CC Sequence 8132 BP; 2248 A; 1977 C; 2037 G; 1870 T; 0 other;

CC Query Match 50.4%; Score 1489.2; DB 22; Length 8132;

CC Best Local Similarity 72.2%; Pred. No. 0;

Matches 2140; Conservative 0; Mismatches 713; Indels 113; Gaps 11;
 QY 1 tgcctttagggtaggaacacccctgagacagtttgggtgacccctatgatatgctc 60
 Db 5257 tgcctttagggtaggaacacccctgagacagtttgggtgacccctatgatatgactc 5316
 QY 61 acgggggaccccccccggtgtgtagaattgtcttcgtatcatagtcgtatgtctctc 120
 Db 5317 acgggggaccccccccggtgtgtagaattgtcttcgtatcatagtcgtatgtctc 5376
 QY 121 cccagccttgtctctccttaagcgaagcgctgagttgtgttgagcaagcggtggaagc 180
 Db 5377 cccagccttgtctctccttaagcgaagcgctgagttgtgttgagcaagcggtggaagc 5436
 QY 181 agctccggagagcctactcaagagagagactgtcaagttccacatcgctcccaagttg 240
 Db 5437 aactccggagagcctactcaagagagagactgtcaagttccacatcgctcccaagttg 5496
 QY 241 gagatcagatctatgttagacgcacacgcgtgcaagaaactcgagactcggtggaagac 300
 Db 5497 gagatcagatctatgttagacgcacacgcgtgcaagaaactcgagactcggtggaagac 5556
 QY 301 ctatctcgtacttttgaccacacacacgcgtgcaagaaactcgagactcggtggaagac 360
 Db 5557 ctatctcgtacttttgaccacacacacgcgtgcaagaaactcgagactcggtggaagac 5616
 QY 361 tccatgcatcccaagctlaagctgctgcacacccctccgactcggtgtgagagcggaaga 420
 Db 5617 tccatgcatcccaagctlaagctgctgcacacccctccgactcggtgtgagagcggaaga 5676
 QY 421 ctgagatcccttaagcttgcctcgcacgtgcgtgtcttctactcctaacaatactcc 480
 Db 5677 ctgagatcccttaagcttgcctcgcacgtgcgtgtcttctactcctaacaatactcc 5736
 QY 481 cagcgcaagttagtaaacgcttlaagacagctcgcacaccccaataagcttccactacc 540
 Db 5737 cagcgcaagttagtaaacgcttlaagacagctcgcacaccccaataagcttccactacc 5796
 QY 541 tggctgattatgacccctgatacgggtgtcaactgtaaaatagcaactcgagtggtctcct 600
 Db 5797 tggctgattatgacccctgatacgggtgtcaactgtaaaatagcaactcgagtggtctcct 5856
 QY 601 agagcacctgtgtgctcgaactgcatcttcgctccgactgattaaaccccgctgtaaa 660
 Db 5857 ttaggaacactgtgtgctcgaactgcatcttcgctccgactgattaaaccccgctgtaaa 5913
 QY 661 agcacactcccaactagtcgtgatagtatggttctatgtctgcacag--cacagag 717
 Db 5914 acctacccccagatactcctcatgctcagagatttattgttgcacagacacacaaat 5973
 QY 718 aagagaaactgtggtggtgttcgtggaaloccttctgttagagatagagctgtcctacc 777
 Db 5974 aatggaacaactgtggtggtgttcgtggaaloccttctgttagagatagagctgtcctacc 6033
 QY 778 tccaacgatagtgagactggaatggtgcgaactctctccagagacgggtlaaattctcctt 837
 Db 6034 tccaacgatagtgagactggaatggtgcgaactctctccagagacgggtlaaattctcctt 6093
 QY 838 gtcaa-----ttccggccgggcaagatgacaaatgatgaaactatataagat 885
 Db 6094 gtcaaactccataccagctctggaacaatttactcctgactgattatgaaactggaagc 6153
 QY 886 aagactctctcccaactgagactgattatgaaatgaaatgaaatgaaatgaaatgaaatg 945
 Db 6154 cccaagtgtctccctccagactgattatgaaatgaaatgaaatgaaatgaaatgaaatg 6213
 QY 946 caggaataatccaagaatgataatgataatgataatgataatgataatgataatgataatg 1002
 Db 6214 caagaataatccaagaatgataatgataatgataatgataatgataatgataatgataatg 6273
 QY 1003 tatggcgggagagcaaggttccacttacaacttgcctttagatagagacgggagcaaga 1062
 Db 6274 tcggtaaacacaacagsgctccatcttaactatctgcctccaataaaccagctg---gag 6330

PD 30-NOV-2000.
 XX 24-MAY-2000; 2000MO-US14296.
 XX 24-MAY-1999; 99US-0135631.
 XX (MAYO-) MAYO MEDICAL VENTURES.
 XX Federpsiel MJ;
 XX WPI; 2001-032041/04.
 DR
 XX Inhibiting or preventing infectious agent transmission in mammalian
 PT transplant recipients, by introducing recombinant DNA comprising DNA
 PT encoding extracellular proteins of the agent into donor cells, such as
 PT swine cells -
 XX
 PS Claim 16; Page 112-115; 144pp; English.
 CC The present invention provides a method to prevent the transmission of
 CC infectious agents during xenotransplantation. This involves introducing
 CC to donor swine cells a recombinant DNA encoding a peptide fragment from
 CC the infectious agent, and then introducing these cells into the
 CC transplant recipient.
 XX
 SQ Sequence 7873 BP; 2200 A; 1913 C; 1962 G; 1798 T; 0 other;
 Query Match 48.5%; Score 1433.8; DB 22; Length 7873;
 Best Local Similarity 70.7%; Pred. NO. 0;
 Matches 2253; Conservative 0; Mismatches 607; Indels 325; Gaps 13;
 QY 1 tgccttttaggttagaagacaccccttgagacagtttggtgctgacccctataatgctt 60
 Db 4649 tgccttttaggttagaagacaccccttgagacagtttggtgctgacccctataatgctt 4708
 QY 61 acggggagaccccccgttgtagaagattgctctgtacatagtgctgagtgctctt 120
 Db 4709 acggggagaccccccgttgtagaagattgctctgtacatagtgctgagtgctctt 4768
 QY 121 cccagccttgctctagagtcagagcgtcgtgagtggtgaggaacagcgctgagagc 180
 Db 4769 cccagccttgctctagagtcagagcgtcgtgagtggtgaggaacagcgctgagagc 4828
 QY 181 agctccggagagcctactcagagagagagagctgtcagatccacatcgcttccagtg 240
 Db 4829 aactccggagagcctactcagagagagagagctgtcagatccacatcgcttccagtg 4888
 QY 241 gagattcagctatagtagagcagcagctgagagaacccctcgagacccgtgtgagagac 300
 Db 4889 gagattcagctatagtagagcagcagctgagagaacccctcgagacccgtgtgagagac 4948
 QY 301 ctctctcgtactcttgacacacacagcgctgtgaaagtccgaagagatccaccctgga 360
 Db 4949 ctctctcgtactcttgacacacacagcgctgtgaaagtccgaagagatccaccctgga 5008
 QY 361 tccatgacacacagcttgaagcttgagcagacccctcgagacccgtgtgagagacgaaaga 420
 Db 5009 tccatgacacacagcttgaagcttgagcagacccctcgagacccgtgtgagagacgaaaga 5068
 QY 421 ctgagaatcccttaagcttgcctcagacagcgtgtgttccctacttaacaataactccc 480
 Db 5069 ctgagaatcccttaagcttgcctcagacagcgtgtgttccctacttaacaataactccc 5128
 QY 481 cagagccagtgtagagcgttatagacagctcgagaccccatagacttataccctacc 540
 Db 5129 cagagccagtgtagagcgttatagacagctcgagaccccatagacttataccctacc 5188
 QY 541 tggctgtatagacccctgagagctgacagcttaaatagacagctgagtgctccct 600
 Db 5189 tggctgtatagacccctgagagctgacagcttaaatagacagctgagtgctccct 5248
 QY 601 agagcaccctgtgtgctgacagctgacattctgctccgattgattaaaccccgctgta-- 658

Db 5249 ttgggagccctgtgagcctgattatatagtctccttcagacgaatccctgtccaat 5308
 QY 659 ----aaagacacccctccacacccctagtcgtagtttaggtttattgtgcctcaag----- 711
 Db 5309 gaccagccacacacccctgtagccgctgcttagaggttttaagcttgcctgagacc 5368
 QY 712 acagagaaagaaatactgtgaggtgtcttgaggaatccctctgtatagagaagagagctgc 771
 Db 5369 ccaataatgaaagaatattgtggaatccctcagagattcttcttgcgaagaatgagagctgc 5428
 QY 772 gtcaactccaagatgagagctggaatgtgagcagatctctccagagacccgtgtaaatc 831
 Db 5429 gtcaactccaagatgagagctggaatgtgagcagatctctccagagacagagtaatc 5488
 QY 832 tccctgtcaatcc----- 846
 Db 5489 tccctgtcaatcc----- 5548
 QY 847 ----ggccggggaagtagaataatgtagaactataaagaataagagctgtcccatc 902
 Db 5549 gattggcaacaggggtacataaagaagtg-tacgaataagcaataaagctgtcattcgt 5607
 QY 903 agacttagattatcctaagaatagttcactgaaagaaagaaacaggaataatccaaga 962
 Db 5608 agacttagattatcctaagaatagttcactgaaagaaagaaacaggaataatccaaga 5667
 QY 963 gtgataaattgtagatagcgtggaatgattttta-----ataatgagcggggagcgag 1019
 Db 5668 gtgataaattgtagatagcgtggaatgattttta-----ataatgagcggggagcgag 5727
 QY 1020 gtcaacttaacacatcgcttagagtagagacgggagacagacccctgtgtgacagtg 1079
 Db 5728 atctgttcgactattcgcttagagtagagacagacccctgtgtgacagtg 5787
 QY 1080 acccgataagtagtctgtgagacaggggccccggccctggagccacagcagataactgac 1139
 Db 5788 acccgataagtagtctgtgagacaggggccccggccctggagccacagcagataactgac 5829
 QY 1140 gttgcccccaatlaactcgtcgctgagacataacacagcgccctagacagtagtccac 1199
 Db 5830 -----acagagccatctccttaaccccttgatttaacaataacaccc 5871
 QY 1200 tggattgattcctacacacacagcctagaaactcccaagtgctcgtttaagacagaga 1259
 Db 5872 tggattgattcctacacacacagcctagaaactcccaagtgctcgtttaagacagaga 5916
 QY 1260 gagactctcagctatcctacagagccttccaagcactcaactccacagcagcagatgc 1319
 Db 5917 gagactctcagctatcctacagagccttccaagcactcaactccacagcagcagatgc 5976
 QY 1320 caactctctgtgtgcttctcctacacagagcctccttcttcttgaagggatgagcaca 1379
 Db 5977 caactctctgtgtgcttctcctacacagagcctccttcttcttgaagggatgagcaca 6036
 QY 1440 taagcttaacccctcagagattccgggagagggagacatgacataggaagaaagctcccccac 1499
 Db 6037 taagcttaacccctcagagattccgggagagggagacatgacataggaagaaagctcccccac 6096
 QY 6097 taagcttaacccctcagagattccgggagagggagacatgacataggaagaaagctcccccac 6156
 Db 1500 ccaagcagaccccttgctatagcagtggttttaagagcagcctcagaaatccaatatt 1559
 QY 6157 ccaagcagaccccttgctatagcagtggttttaagagcagcctcagaaatccaatatt 6216
 Db 1560 agtaccctgttataacaggtgtgtgagcagacatctgtgttaacccctgtgttccac 1619
 QY 6217 ggtaccctgttataacaggtgtgtgagcagacatctgtgttaacccctgtgttccac 6276
 Db 1620 ctgagcttcaacccaatccaagattgtgtgtcagtggtccaatcgttcccgagtgta 1679

QY	2596	-----gcaccatagaagaattgattacacattgacagccctag	2633
Db	7406	gagdgagaagctcaattgccttaacggaacccacgtagatcgggtgttaccacaaaagtgtgaa	7465
QY	2634	tgacctactcaactgtaacatctgtcaactctgcgccaaagagccacagcagatgcygactcc	2693
Db	7466	acacatactcttggacacacatgtctctcccccaccgcgaacatgycgaatgtgttaactc	7525
QY	2654	ggagcatttttaaaatgatttgggtgccacgagcgcggtctcgcataattttaaatgattg	2733
Db	7526	taaaacattttaaatattatgtgtccacgaagcgcggtcctcgaagtlttaattgactg	7585
QY	2754	gtccatgtgagcgcggtcctcgcataattttaaaatgattgtttgttgaagcacaagcttg	2813
Db	7566	gt-----tltgtgatactttgaaatgattggtttgtaaagcgcgggtcttg	7630
QY	2814	tgtgtaaccccatataaagctgtgtccgattccgcactcggcgccagctccttaacctg	2873
Db	7631	ttgttaaaccccatataaagctgtgtccgactcgcacaccccggtgcgcagctccttaacctg	7690
QY	2874	cgtgtgttaacgactgtgtggccccagcgcgcttggaaataaaatcctctgtctgttgc	2933
Db	7691	cgltgtgttaacgactgtgtggccccagcgcgcttggaaataaaatcctctgtctgttgc	7750
QY	2934	caaaa	2938
Db	7751	caaga	7755
RESULT 11			
ID	AAT74884		
XX	AAT74884	standard; cDNA; 7892 BP.	
XX	AAT74884:		
DT	09-FEB-1998	(first entry)	
XX			
DE	Miniature swine retrovirus cDNA.		
KW	Retrovirus; porcine; GAG protein; POL protein; ENV protein;		
KW	xenotransplantation; infectious; provirus; organ transplant; donor;		
KW	activated virus; PCR; ss.		
OS	Porcine retrovirus.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	585..2159	
FT		/*tag= a	
FT	mat_peptide	585..2156	
FT		/*tag= b	
FT	CDS	/note= "putative GAG protein"	
FT		2307..5744	
FT		/*tag= c	
FT	mat_peptide	2307..5741	
FT		/*tag= d	
FT	CDS	5620..7536	
FT		/*tag= e	
FT	mat_peptide	5620..7533	
FT		/*tag= f	
FT		/note= "putative ENV protein"	
PN	W09721836-A1.		
PD	19-JUN-1997.		
XX			
XX	13-DEC-1996;	96WO-US19680.	
XX			
PR	14-DEC-1995;	95US-0572645.	
PA	(GEHO) GEN HOSPITAL CORP.		
XX			

PI Fishman JA;

AA WPI; 1997-332804/30
DR

DR P-PSDB; AAW32096-W32098.

PT	New nucleic acid from porcine retro:viruses - used for detecting
PT	viruses in transplant or other tissue and for assessing risk of
PT	transmitting infection to graft recipient

PS Claim 22; Fig 3; 128pp; English.

This cDNA sequence represents a porcine retrovirus from miniature swine containing the putative coding regions for viral GAG, POL and ENV proteins. This sequence and PCR fragments generated from it (see AAT74812-74882) can be used to screen organs for the presence of porcine retroviruses prior to xenotransplantation. Transplantation can increase the likelihood of retroviral activation if intact and infectious proviruses are present. The porcine retroviral sequence can be used to generate probes to determine the level (e.g. copy number) of intact (i.e. potentially replicating) porcine provirus sequences in a strain of xenograft transplantation donors. It can be used to detect mutations, genetic lesions or viral recombinants and also to determine the histological localisation of activated retroviruses. Using Polymerase Chain Reaction DNA Quantitation (PQO) on blood mononuclear cells, infectivity titration and susceptibility testing can be performed. Ultimately animal donors without intact porcine retroviral sequences or a lower copy number of viral elements could be selected.

SQ Sequence 7892 BP; 2191 A; 1915 C; 1980 G; 1806 T; 0 other;

Query Match	45.6%;	Score 1347.8;	DB 18;	Length 7892;
Best Local Similarity	74.9%;	Pred. No. 0;		
Matches 1834;	Conservative	0;	Mismatches 522;	Indels 93;
			Gaps	8

QY	1	tgctcttaagggttaggaacacccctgagcagtttgggctgcagcccccctatagtaatgctct	60
Db	5257	tgctcttttgggttttaggaacacccctgagcagtttgggctgcagcccccctatagtaatgctct	5316
QY	61	acggggggaaccccccccggttggtagaatgtctctgtacatagtctgtagtgcgtctt	120
Db	5317	acggggggaaccccccccggttggtagaatgtctctgtacatagtctgtagtgcgtctt	5376
QY	121	cccaagcctttgtctctctagagctcaagcgctcagatgggttgagcaacagcgttgaagc	180
Db	5377	cccaagcctttgtctctctagagctcaagcgctcagatgggttgagcaacagcgttgaagc	5436
QY	181	agcttcggagagcctactatagagagagagacttgcgaatctcaactgccttccaagt	240
Db	5437	aacttcggagagcctactatagagagagagacttgcgaatctcaactgccttccaagt	5496
QY	241	gagatcagatctatgtttagacgcacacgcgtgcacgagaaacctgcagactcgttgaagagac	300
Db	5497	ggagttcagttcaaggttagacgcacacgcgtgcacgagaaacctgcagactcgttgaagagac	5556
QY	301	cttatctcgtaacttttgaccacacccaagcgctgttgaagtcgaaggaatctccaactgga	360
Db	5557	cttatctcgtaacttttgaccacacccaagcgctgttgaagtcgaaggaatctccaactgga	5616
QY	361	tccatgcatacccaagtttaagcttggcgccacccctccccaactcgggttgaggaagccgaaaga	420
Db	5617	tccatgcatacccaagtttaagcttggcgccacccctccccaactcgggttgaggaagccgaaaga	5676
QY	421	ctgagaatcccttaagcttcgctccatcgccttggtctcctactactaacaataactcc	480
Db	5677	ctgagaatcccttaagcttcgctccatcgccttggtctcctactactaacaataactct	5736
QY	481	cagggcagagtaaaagccttatagacacgtcgaaaccccatagaccttatcccttacc	540
Db	5737	caggacaaatggtatagcctataggaagcagccttgaaactcccaataaaccttactccacc	5796
QY	541	tggctgattatttgaccttatacgggttgcactgttaaatagactcagagtttgctctct	600

D	b	5797	tggttaattactgctatccgycagaggttataataatacaaacactccaaggggaggtctct	5856
Q	y	601	aagagcacttggctggccttgacatgcattcttcgctccgcattgataaccgcgtgttaa	660
D	b	5857	tttagaacctcgggtggccttgatctatacagtttggcttcagatccagttatctc---ctagctcg	5913
Q	y	661	agcacactcccaaccctagtcggtagttaatgtgttattgtgtcccgagccag---ag	717
D	b	5914	aacctaccctccagatataccctccatctgcacagattttaaagtttgcgcagagaccacaaat	5973
Q	y	718	aaagagaatactctggtgggtctctcgggaatccctctctgttagagatagggcgtccac	777
D	b	5974	aatggaaaaacttgcgggaataatccagagattctctcttcttgaacaatagcaacgtgtaac	6033
Q	y	778	tccaacgatgtgagacttggaaatctgcgcatctctccacagaccgggtlaaattctctctt	837
D	b	6034	tctaagtatgatataatgtgaatctgccaacctccagcagatagagtaagtttctctat	6093
Q	y	838	gtcaa-----ttccggccggcggaagtatacaaatgatgaactatataaagt	885
D	b	6094	gtcacacctatacacaagctcttgcagaaatttaattaccttgcaccttgattagaacctgaa	6155
Q	y	886	aagagcttgcctcccatcagactcttagattatctaaagataagtttcaacttgaagaagaa	945
D	b	6154	ccaagtgctctctccatcgaagcactagattacctaataataaagtttcaacttgaagaagaa	6213
Q	y	946	caggaaataattccaagaagtgtgataaatgtgatagcttgggagataagtttta---ata	10023
D	b	6214	caagaaataattaccctaataatgtgtaaatgtatgtctcttgggagatgtatatattagagc	6273
Q	y	1003	tatggcggggagagaggttccactttaaccattgcgcttagatagagacgggacagaa	1065
D	b	6274	tcgggtataacaacacagagctctcatcttaactatctgcctcaataataaacacagctg--	929
Q	y	1063	cccccttgggaataaggagaccggcctaaagtatctgcgcgaacaggggcggccgccttgag	11222
D	b	6331	ctcccaatggtcataagagcaataatagcgtcttgacgggtccaagaaccccaacccaagaa	6390
Q	y	1123	ccacgcgataactctggccgggtgccccaattaacctctgtcgggcgttgcataaacagacgg	11825
D	b	6391	ccaagacatctctt-----	6405
Q	y	1183	ccctagcaaacagttacacttgatttgattcctaccacaacgcttagaaactccccagtgct	1244
D	b	6406	-----aacataaactcttgatctcagacccacatgagttctaaacagccagac-----	6449
Q	y	1243	ccctttagaacagagacagagagactcttcagctcatcccaaggagccttccaagcactaac	13023
D	b	6450	-----taaaatgggggccaanaacttttagaccctaccccaaggagccttttaagctctaac	6504
Q	y	1303	tccaaccgacccctgtatgcacacttctctctgtctgttgcattcctcaagggcctctctat	13625
D	b	6505	tccaagactccaagaggtctactcctctctctgtgtctatgtcttagcttcgggccacctaac	6566
Q	y	1363	tatgaggggattgcttaagaagaagaattcaatgtgacacaagagcatagaatacaatgt	14223
D	b	6565	tatgagaagaatgcttagaagaagaaattcaatgtgacacaagaacatagagacaactatgc	6624
Q	y	1423	acatgtgggtgtcccgaaataagcttaccctacacgaagtttccgggaaagggcagcatata	1484
D	b	6625	acatgtgggtatcccaataataagcttacccttactcgaagtttcttggaaaaaggccctgata	6684
Q	y	1483	ggaagaagcttccccaatcccaacaacacttctgtatagtaactgtgttattagacagcc	15443
D	b	6665	ggaaaaggttcccccatctcccaacaacaccttctgtaaccacacatgaaagcctttaacaac	6744
Q	y	1543	tcaagaaatcagattttagtaacctgtgtataacaggttgggtggacatgcaatactgggta	1607
D	b	6745	tcctgagatccaatatctgttaacctgtgtatagacaggttgggtgacatgataatcgtatata	6804
Q	y	1603	acccccctgttttccacctcaagcttctaacaacatccaagaagtgtgtatgttgtcaaa	1665
D	b	6805	acccccctgttttccacctctgtttttaacaaaactaaagaatttttggatatgtgtccaa	6866

FT	/note= "putative ENV protein"
FT	3112..4686
FT	/*tag= b
FT	/note= "putative GAG protein"
FT	4871..8060
FT	CDS
FT	/tag= c
FT	/note= "putative POL protein (partial)"
XX	
PN	MO9721836-A1.
XX	
PD	19-JUN-1997.
XX	
PF	13-DEC-1996; 96WO-US19680.
XX	
PR	14-DEC-1995; 95US-0572645.
XX	
PA	(GEHO) GEN HOSPITAL CORP.
XX	
PI	Fishman JA;
DR	WPI; 1997-332804/30.
XX	
PT	New nucleic acid from porcine retroviruses - used for detecting
PT	viruses in transplant or other tissue and for assessing risk of
PT	transmitting infection to graft recipient
XX	
PS	Claim 1; Fig 1; 128pp; English.
CC	This sequence represents the purified porcine retroviral cDNA
CC	sequence of Tsukuba-1 and contains the putative coding regions for viral
CC	proteins GAG, POL and ENV. This sequence and PCR fragments generated
CC	from the sequence (see AAT74812-T/4882) could be used to screen organs
CC	for porcine retroviruses prior to xenotransplantation. Transplantation
CC	can increase the likelihood of retroviral activation if intact and
CC	infectious proviruses are present. The porcine retroviral sequence can be
CC	used to generate probes to determine the level (e.g. copy number) of
CC	intact (i.e. potentially replicating) porcine provirus sequences in a
CC	strain of xenograft transplantation donors. It can be used to detect
CC	mutations, genetic lesions or viral recombinants and to determine the
CC	histological localisation of activated retrovirus. Using Polymerase Chain
CC	Reaction DNA Quantitation (PQO) on blood mononuclear cells, infectivity
CC	variation and susceptibility testing can be performed. Ultimately animal
CC	donors without intact porcine retroviral sequences or with a lower copy
CC	number of viral elements could be selected.
XX	
SQ	Sequence 8060 BP; 2233 A; 1959 C; 2012 G; 1856 T; 0 other;
<hr/>	
Query Match	41.7%; Score 1231.8; DB 18; Length 8060;
Best Local Similarity	69.8%; Pred No. 0;
Matches 1873; Conservative	0; Mismatches 697; Indels 113; Gaps 11
QY	279 ctcgagacctcgtgtaagaaggacctatcttcgtactttgaccacaaccaagctgtgaaa 338
b	1 ctcgagactcgtgtaagaaggaccttattctgcattttgaccacaaccaagctgtgaaa 60
QY	339 gtccgaagaagtcccacccctggatcatgatcatcccaacgttaagctggcgccaacctccgac 398
D	61 gtccgaagaagtcccacccctggatcatgatcatcccaacgttaagctggcgccaacctccgac 120
QY	399 tcgggtggtggaagccgaagaagctagatacccccttaagcttcggccatcgctggt 458
D	121 tcgggtggtggaagccgaagaagctagatacccccttaagcttcggccatcgctggt 180
QY	459 ccttactcttaacaaataactcccccaagcgacgtatgaacagccttatacacagctgaacc 518
D	181 ccttactcttctcaataaactctcaacataatgtagtgcatagagaagacgctgaactc 240
QY	519 ccataagaccttacccttactgtgctatattgtaaaccttcggtatcggtttaa 578
D	241 ccataaacccttatctctcaactcgtttaataatcagaccgccagagctaataatcaa 300
Y	579 taagactcgaagtgcttctcctaagaggaacctgtgtgcttgaactgtgattctgcttcg 638

QY 2776 atattttaaattgtgtttgtgacgacagcttctgttgtaacccataaagctgt 2835
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 2431 atatttgaattgtgtttgttgaagcggtgttctgttgtaacccataaagctgt 2490
QY 2836 ccgacttcgcactcgcggcgccgactctactacccctgctgtgtgtacgaactgtggccc 2895
||||| ||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 2491 ccgacttcacactcgcggcgccgactctactacccctgctgtgtgtacgaactgtggccc 2550
QY 2896 cagcgctgtggaataaaacccctctgctgtgttgatcaaaa 2938
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 2551 cagcgctgtggaataaaacccctctgctgttgatcaaaa 2593

RESULT 13

AA77725
ID AAF77725 standard; cDNA; 8060 BP.

AC AAF77725;

DT 23-MAY-2001 (first entry)

XX Tsukuba-1 cDNA.

XX Tsukuba-1; retrovirus; graft transplantation; xenotransplantation; ss.

OS Porcine retrovirus.

XX US6190861-B1.

PN 20-FEB-2001.

XX 13-DEC-1996; 96US-0766528.

XX 14-DEC-1995; 95US-0572645.

PR (GEHO) GEN HOSPITAL CORP.

XX Fishman JA;

XX WPI; 2001-256211/26.

PT Assessing risk of endogenous retroviruses in clinical practice and in
PT xenotransplantation, comprises using probe sequences derived from swine
PT or miniature swine retroviral genome -

PS Claim 1; Fig 1; 127pp; English.

CC The present invention relates to a method for screening a cell or tissue
CC for the presence or expression of a retrovirus (RV), comprising
CC contacting a target nucleic acid from the cell or tissue with a second
CC nucleic acid from the present invention (e.g. the present sequence or a
CC fragment thereof). The method is useful for RV detection and to assess
CC graft transplantation risk. Screening of animals allows the elimination
CC of donors with active replication of known viruses. Inactive proviruses
CC can be detected and inactivated, allowing identification and elimination
CC of potential human pathogens derived from swine in a manner not possible
CC in the outbred human organ donor population and is important to the
CC development of human xenotransplantation.

CC Sequence 8060 BP; 2233 A; 1959 C; 2012 G; 1856 T; 0 other;

Query Match

Best Local Similarity 41.7%; Score 1231.8; DB 22; Length 8060;
Matches 1873; Conservative 0; Mismatches 697; Indels 113; Gaps 11;

QY 279 ctgagactcgtgtgaaggacctatctcgtactttgacacacacacgctgtgaaa 338
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1 ctgagactcgtgtgaaggacctatctcgtactttgacacacacacgctgtgaaa 60
QY 339 gtggaaggaatctccacccgtgactcatcccaagcttaagctgtggccacctccgac 398
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 61 gtccgaaggaatctccacccgtgactcatgcatccccacacgttaagccggcgccacctccgac 120
QY 399 tccgggttggaagacggaagactggaatcccttaagcttcgcctccatgcctgtgt 458
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 121 tccgggttggaagacggaagactggaatcccttaagcttcgcctccatgcctgtgtgt 180
QY 459 ccttacttaacataactcccccagtcagtgaaagccttatagacagctcgaaacc 518
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 181 ccttacttcgtcaataactcctcagactaatgtgatgacatagagaagacgtgaacc 240
QY 519 ccatgaaccttatccctacactcgtgctgtatattgaacctgatacaggtgtgactgtaaa 578
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||||
Db 241 ccatgaaccttatccctacactcgtgctgtatattgaacctgatacaggtgtgactgtaaa 300
QY 579 tagcactcgaaggtgtgtccctcctagaagcactggtgtgctgaactgtatctgcctccg 638
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||||
Db 301 caacactcaagggaggtccctcttgaagaaacctgtgtgctgtactataagtttgcctccag 360
QY 639 atgtatgaaccccgctgttaaaagcacacctcccaacctagtcctgtatgtgtgttca 698
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 atcagttatc---ctagcttgacctcaccccccagatatccctccatgctcagagatttta 417
QY 699 ttgctgtcccaag---cacagagaagaagaatactgtgtgggtgtctgtggaaacctctgt 755
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 418 ttgtgtcccaagaccccaataatgaaacattgtgcgaataatcccaagatttcttltgt 477
QY 756 taggaatgtgagctgtgctactcccaagatggaactgtgaaatgtgcgactctctcca 815
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 478 taacaatgtgaactgtgtactcctcattatgtatattgtgaatgtgccaaacctccaca 537
QY 816 ggaacggtaaaatttcctctgtcaa-----ttccggccggggaagttaaa 863
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 538 ggaatgaggaagtgttcttctatgtcaacacctataccagctctgtgacaatttaacct 597
QY 864 aatgtatgaactatataaagaataagctgtcctccatcaagactgaattatcctaagat 923
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 598 gacctgtattgaactgtgaagcaccaggtgctcctcctcctgaagcactgtattactaaaat 657
QY 924 aagtttcaactgaaaaaggaagaaacaggaatattcaaaagtgtgataatgtatgaactgt 983
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 658 aagtttcaactgaaaaaggaagaaacaggaatattcaaaagtgtgataatgtatgaactgt 717
QY 984 ggaatagttttta---ataalalgcygggaggaaggtgtcacttcaaacactgcct 1040
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 718 ggaataggtatatttgtaggtcgtcggtgaaacacacaggtctcatctcaactatctgcct 777
QY 1041 taggatagagacggtgacagaaacccctgtgtgcaatgtgaaacccgataagactgtgctga 1100
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 778 caaataaacacagctg---gagcctccaaatgtcctaagaccaaatacgttctgaagg 834
QY 1101 acaggggcccccgccctcggaagccacgataactgtccggtgtcccaattaaactgtct 1160
||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 835 tcaaaagacccccaacccaagagacacacacactcct----- 871
QY 1161 gggcgtgacataacacagcgcctcagacacagtcacactgattgtcttaccacaacac 1220
||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 872 -----aacataactctgtatagaagcccaactgagtc 903
QY 1221 gcttgaagaactcccaagtgctcctgttaagaagagacagactctcagctcatcca 1280
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 904 tagcagca-----cgactaaatagggggcaaaacttttagcctcatcca 948
QY 1281 gggagcttccaagcactcaactccaaccccgatgtccaactcttctgtgtgtgt 1340
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 949 gggagcttccaagcactcaactccaaccccgatgtccaactcttctgtgtgtgt 1008
QY 1341 tcatcctcaaggcctcctctatattgaagggtgtgctaaagaagaagaatcaatgtgac 1400
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1009 cttagcttggcccaactctactaagaagatgtgctagaagagggaatcaatgtgac 1068
QY 1401 caaagaacatagaataatgaatgtgaggtgtcccgaaataagcttaccactgaagt 1460
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1069 aaaagaacatagaacaaatgcacatggggtgacccaaataaagcttaccactgaagt 1128

QY	1461	ttccgggaaggaggccatgcctatagaagaagctctcccatccacccacacacttgcctatg	1520
Db	1129	tttcggaaaaggccaccgcctatgaaaggtctcccatccacacacacttgcctatg	1188
QY	1521	tactgtgtttaatgagcagcgcctcagaaaatcagtaattagtaactcgtgtataacagtg	1580
Db	1189	cactgaagcccttaatacaaacctcgaagtccaatctcgttaactcgtgtatgacagtg	1248
QY	1581	gtggagcgtgcatactgtggtttaacccctgtgtttccacctcagttcttaaccatcca	1648
Db	1249	gtggagcagtaatactcgtatlaacccctgtgtttccacctcgtgtttttaaccaactaa	1308
QY	1641	agattctgtgtcattgttccaacatcgtcccccagatgtatcaaccctcgtgaagttgt	1700
Db	1309	agattttgcatatgttccaacatgtctcccgagtgatatactatcccgaaaagaagcat	1368
QY	1701	ccctgtatgatalatgatactatcgttataacccgacacaaagaagaccgtatccctaac	1760
Db	1369	ccctgtatgatalatgatactatcagtaacaaagaagaccatctctctgaac	1428
QY	1761	agcgttaatgtctgattgaagcgcgcgttgcgtatgaagacagggagacagctgcctat	1820
Db	1429	tgctgtgatgtcgtgaccttggagttggaagcaggtgtgaaagacagacagcgtccctgt	1488
QY	1821	cacagagcaacagcagctatagaagaagacttctgtgacatactgcgcacatgacataga	1880
Db	1489	cacggagcaacagcagctatagaacaagactttagtaacctatcagatgtgtaacagaaga	1548
QY	1881	ttctccgagccttaagagagctgtgttgcaaccacagaagaagtcctcgtaccttctgtcga	1940
Db	1549	ttctcccaagcccttaaaaaaactctgcagtaacctcggaggaatcccttaacctctatcga	1608
QY	1941	agtggtttctaagacacccggagggagatagatactctgcgtttcttaagagaaggtgcgtatg	2000
Db	1609	agtagtccctacagataagaaggggttagatttatattcttaaaagaaggaggtatg	1668
QY	2001	tgcaagccttaaaaagaagaatgttgcctcttatgtgaatcaactcagaagccataagagact	2060
Db	1669	tgtagccttgaaggaagaatgctgtttttagtgtgatactacaaggccatcagaagact	1728
QY	2061	catbaacaaagctctgaaaaaaagttagaagcgtctgaagggagaaggaagcgtgacaaagg	2120
Db	1729	catgaacaacacttagaagaagaaggttggagaagcgtctgaagggagaaggaacataccaagg	1788
QY	2121	gtggtttgaagaatggtttcaacaggtctcctgtgatgacacacccgtcttctgtcctgac	2180
Db	1789	gtgttttgaaggagatgtgttcaacaggtctcctgtgtgtgactacacttctgtcttaac	1848
QY	2181	gggggcacctagtaactcgtctcgttacttaagaattgggccttgccttaataatagtt	2240
Db	1849	agagcccttaataagctccctccctcgttaaccacagaatctgggccaatgatataatacaagt	1908
QY	2241	tgtagccttgtttgagaagaagatgtgattgcagtcacagacatcgttacttaggcaacagta	2300
Db	1909	aattgccttcaattagagaacgataatgtagtgcagatcatcgttacttagcaacagta	1968
QY	2301	ccaaagcctcttgagcccaagagaagaactgacctctagccttccagttccaaatgaatga	2360
Db	1969	ccaaagcccgctctgac--aggaagcctggccgcttagctctacacagttctaagattga	2028
QY	2361	ctattacaagaacaaagatctggggaatgtgaagatbaaatagtacaacthaacccctccga	2420
Db	2026	ctattacaagaagagaagatcgggggaatgtgaagaagatgaataatacacttaagtaattgga	2088
QY	2421	accacaggaagttaataaaaaagctcaaatgcccccgaaattacagacccctgtcgtgcga	2480
Db	2086	agcttaaatatgtctctaattccagaagttgttcccttaatagtaaaagaattagttttt	2148
QY	2481	gtaaatgtgtagaaggtcacaaactcctatgtgtccagggccctgcctatcctgacct-aagt	2539
Db	2146	gcgtgttcaaatatcgggaagtaaaaatggccctgtgataactgtctctcagatgcaagaac	2208

XX 05-JUN-2000 (first entry)
 XX Contiguous DNA of porcine endogenous retrovirus-D (PERV-D) env region.
 DE Porcine endogenous retrovirus; PERV-D; virucide; prevention; vaccine;
 KW pig; diagnosis; infection; xenotransplantation; antibody; env region; ds.
 XX Sus scrofa.
 OS
 XX MO20001187-A1.
 PN
 XX 02-MAR-2000.
 PD
 XX 18-AUG-1999; 99WO-US19053.
 PF
 XX 18-AUG-1998; 98US-0097015.
 PR
 XX (BIOT-) BIO TRANSPLANT INC.
 PA
 XX Banerjee PT, Patience C, Andersson GK;
 PI
 XX WPI; 2000-224704/19.
 DR
 XX Porcine retroviral PERV-D polypeptides for diagnosing porcine
 PT retroviral infections in humans after xenotransplantation -
 XX
 XX Example 4; Fig 11; 119pp; English.
 PS
 XX The present DNA is the contiguous sequence of porcine endogenous
 CC retrovirus-D (PERV-D) env region. It is compiled from the DNA fragments
 CC obtained from the 5' end, extended proline rich region and 3' end of
 CC PERV-D env region. It is isolated from the porcine kidney, PK15 cell line
 CC (ATCC No. CCL-33). PERV-D DNA has 79% homology to a portion of PERV-C and
 CC has virucidal activity. PERV-D sequence is useful for prevention or
 CC diagnosis of infection of human tissues by porcine retroviruses after
 CC xenotransplantation procedures. PERV-D polypeptide may be used to produce
 CC specific antibodies, that can be administered as vaccines to create
 CC passive immunity. The DNA can be used as a hybridisation probe or primer
 CC for isolation purposes.
 CC
 CC Sequence 2000 BP; 667 A; 468 C; 366 G; 499 T; 0 other;
 SQ

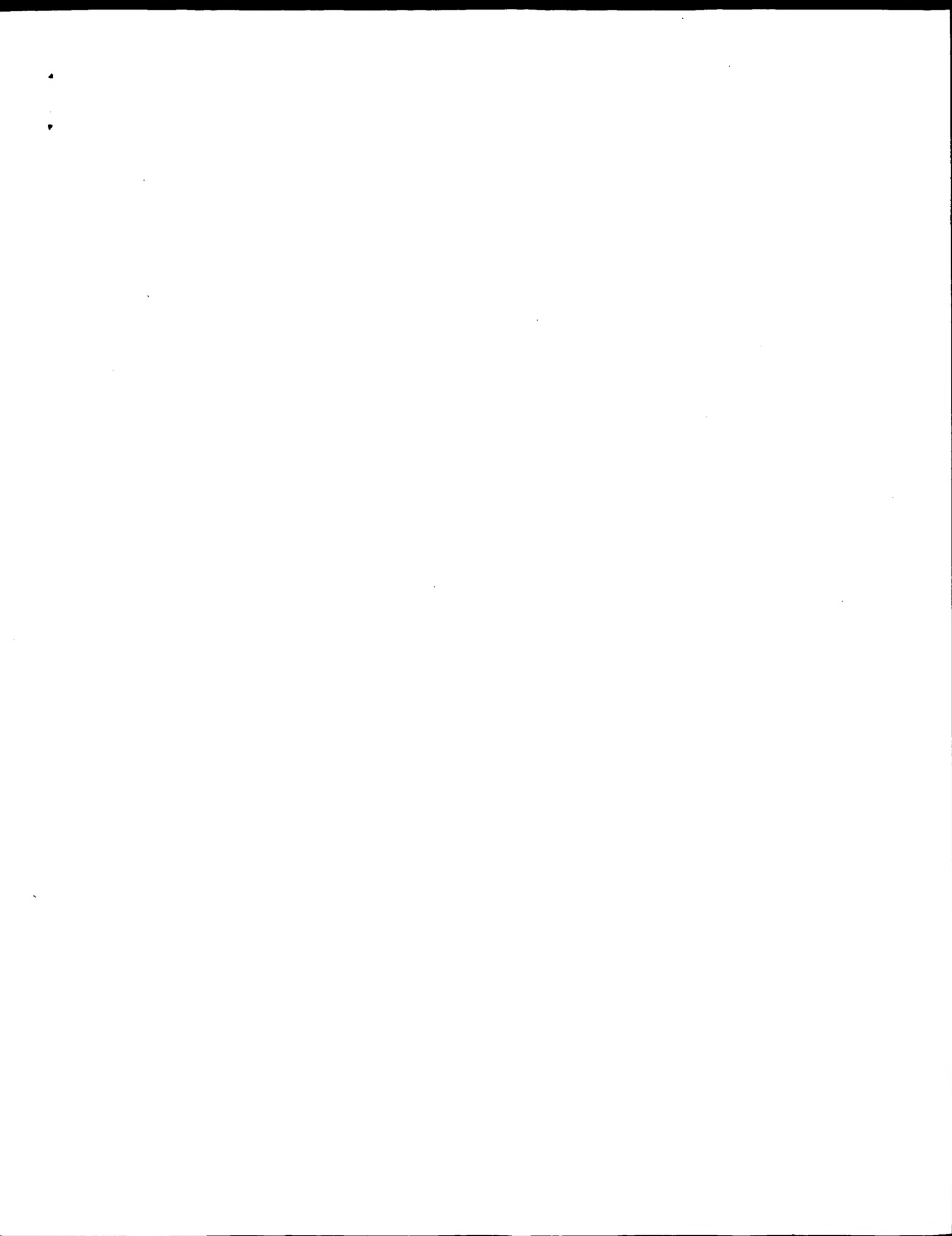
Query Match 32.9%; Score 973.8; DB 21; Length 2000;
 Best Local Similarity 70.2%; Pred. No. 8; 2e-277;
 Matches 1461; Conservative 0; Mismatches 532; Indels 88; Gaps 8;

QY 322 caccacagctgtgaagtggaaggaatccacacgtgacatccacgttaagc 381
 Db 1 caccacagctgtgaagtggaaggaatccacacgtgacatccacgttaagc 60
 QY 382 tggcgcaactcccgacgtcggtgagagccgaagaatggaatcccttaagctc 441
 Db 61 cggcgcaactcccgacgtcggtgagagccgaagaatggaatcccttaagctc 120
 QY 442 gctccatcgctgtgtcttactactcaataaactcccgagccagtagtaagcctt 501
 Db 121 gctccatcgctgtgtcttactactcaataaactcccgagccagtagtaagcctt 180
 QY 502 atgaagcgtcgaacccccaatagacatttccacacgtggtgattatgacacgtat 561
 Db 181 ggaagcgtcgaacccccaatagacatttccacacgtggtgattatgacacgtat 240
 QY 562 acggtgtacactgaataaactgaactgagtggtctcctgaagcactgtgtgctgaa 621
 Db 241 acggtgtacactgaataaactgaactgagtggtctcctgaagcactgtgtgctgaa 300
 QY 622 ctgacattctgctcgatgataaccgcgtgttaaaagcaacactcccaactagtc 681
 Db 301 ctgacattctgctcgatgataaccgcgtgttaaaagcaacactcccaactagtc 357
 QY 682 cgtatgtatggtgtctctatgtctgcccag---cacagagaagaatactgtgtggggt 738

Db 358 cgtgtctacggaatttctgtgtcccgagaccaccaataatggaataacactatggaat 417
 QY 739 tctgggaatccctctctgagagatgagctgcgtccatccacgtatgagacttgaaa 798
 Db 418 cctagaagttctctttaaacaacatgagcgtgttaaccttaatgtatgaaatcgaaaa 477
 QY 799 tggccatctctccagacccggttaaatctctctgttcaatcccgccggcgcaag 858
 Db 478 tggccatctctctcagataggttaagcttctcttcttcaaccccaataacaccog 537
 QY 859 tacaatagatgaactatataaagataagggctgcctcccatcagactatgatacta 918
 Db 538 acctggaatacgtcataggtacatctctgtgttctcccatcagactatgatacta 597
 QY 919 aagataagttctacgtgaataaagaaacaggaataatccaagtgtgataatgtatgt 978
 Db 598 aataatagttctacgtgaataaagaaacaggaataatccaagtgtgataatgtatgt 657
 QY 979 agctgggaataatgttttta---ataatagcgggggagagaggtccacttaacatt 1035
 Db 658 tcttgaggatataatataatatacaggttgagacagacacaggtccatcttaacatt 717
 QY 1036 cgccttagatagaagacgggagacacccctgtgtgcaatgagacccgataatgact 1095
 Db 718 cgaactaaata---agccagtagagccccaatgtctataagacggaataggttcta 774
 QY 1096 gctgaacaggggcccccgccctgtgagcacgcgaatactgtgcggtgcccaatgaac 1155
 Db 775 acgggttaagaagacccaacccagacatctct----- 810
 QY 1156 tgcgtgcggcctgacataaagaagcgcctgtgcaacagtagtaccctgtatcttacc 1215
 Db 811 -----galataatcttaatttgaacccact 837
 QY 1216 aacacgctagaactcccccaggtgttctctgttaagacagagacagactctacgtc 1275
 Db 838 ggtgttaagaagacagac-----taaacggggacaaacttttagtctc 882
 QY 1276 atccagggagcttcccaagcactcaactccacacagcagctgagcactctctgttg 1335
 Db 883 atccagggagcttcccaagcactcaactccacacagcagctgagcactctctgttg 942
 QY 1336 ctgtgtctatctctcagggcctctcttattatgagggagctgagcgaagaagaataat 1395
 Db 943 ctgtgtctatctctcagggcctctcttattatgagggagctgagcgaagaagaataat 1002
 QY 1396 gtgaccaaagacatagaataatcaatgtagtgggtccggaataaagcttaacact 1455
 Db 1003 gtgaccaaagacatagaataatcaatgtagtgggtccggaataaagcttaacact 1062
 QY 1456 gaagttccgggaagggagcatgcataggaagaagctccccaatccacacacacttgc 1515
 Db 1063 gaagttccgggaagggagcatgcataggaagaagctccccaatccacacacacttgc 1122
 QY 1516 tatagactgtgtttatgagcaggtctcagaagaatgtagtattgactcgtgttatac 1575
 Db 1123 aaccacacaggaagcctttaaacaacccctgaggaatcatatctgtgtacccgttat 1182
 QY 1576 aggtgtgggcatgcaatctgtgttaaccccttgttccacactcagcttcaacaa 1635
 Db 1183 aggtgtgggcatgcaatctgtgttaaccccttgttccacactcagcttcaacaa 1242
 QY 1636 tccaaagattgtgtcatgtgtccaatcgctcccgagtgtagtaccatccctgagaa 1695
 Db 1243 actaaagactttacattatgtgtccaatgtgtcccgagtagtaccatccctgagaa 1302
 QY 1696 gtgtgtcctgtatgataatgactatgtagtataacggcaaaaaggaacccgttatcc 1755
 Db 1303 acaattctctgataatgataatgataacggcaaaaaggaacccgttatcc 1362
 QY 1756 accctagctgataatgctcgataggaagcggcgtgtgctgtaggaacagggagcgtg 1815

Db	1363	aaactcgcgaatgctgctgcagctcgcgaagtgataacagctgtagagaacaggactgcagct	1422
QY	1816	ctgatacaggacacacacagcagctcagagaaagcactgctgtagctacatgcgccatgaca	1875
Db	1423	ttagttacagagccttcagcagctcagataacacagcacttagtaacctacatacaaltgtaaca	1482
QY	1876	gaagatctccgaagccttaaaagagctctgttaagcaacctcgaagagtcctgaactctcttg	1935
Db	1483	ggaatctcccaagcctcaaaaaaatctgctcgaagaaactcggaaaaatccctaacctccctta	1542
QY	1936	tcctgaagctggtcttaccagaaacccgagagggattagatctgcttcttctatagaagagctggg	1995
Db	1543	tcctgaagtagtctacagagataaaaaaggttagatttaattcttaaaaaaaagaaga	1602
QY	1996	ttatgctcagccttaaaagaagaatctgctctctatgtatagatcactcagcagagccatcaga	2055
Db	1603	ttatgcttgaagccttaaaagggaaaatgctgttctttagttagatcattcttcaggggcattcaga	1662
QY	2056	gaccctcctgacaagcctctagaaaaaaggtttagaagggcgcgcgaagggaaaagagagctgcac	2115
Db	1663	gaccctcatgacaagcctcttaaaaaaggttctggagaaacgctcgaagggaaaaggaactcac	1722
QY	2116	-caggggctggtcttgaagagatggtctcaacagagctcctcttgaaagacacccctgctctgc	2174
Db	1723	tcaaggtggttttaaaagatggtctcaacagctctcctctggttggtgtacctacttctctac	1782
QY	2175	tcctgaagggcccttagttagtctctgctctctgttacttaccagcttgtagccttgccttaattaa	2234
Db	1783	tttaacaggagcccttaataagctctctcctctctgttaccacaaagtttgggccaagtatattaaaa	1842
QY	2235	taggttctggtccttctttagagaagaagtgtagtgagctcagatcacaatgcttacttaaggca	2294
Db	1843	caagttatctgctctcaatttagaaaagaataagtgcagctccagatcacaatgcttacttagaca	1902
QY	2295	acagttaccaaagcctctcttgagccaagagaaactgaaccttagcctctccacgttctaaga	2354
Db	1903	acagttaccaaaccccatcta---acagggaagctgycgcgttagctctaccagttctaaga	1959
QY	2355	ttagaactattaaacagaacagaagctgggggaattgaagaagt	2395
Db	1960	ttaaacctattaaacagaagaagctgggggaattgaagaagt	2000

Search completed: February 24, 2002, 03:05:27
Job time: 19493 sec



OM of: US-09-171-553b-10 to: N_Geneseq_1101:* out_format : pfs
 Date: Feb 24, 2002 10:24 AM
 About: Results were produced by the GenCore software, version 4.5,
 Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=framet p2n.model -DEV=xlh
 -O=/cgn2.1/USPTO.spool/US09171553/runat_22022002_122826_11203/app_query.fasta_1.3312
 -DB=N_Geneseq_1101 -QFMT=fastcap -SUFFIX=feb22std.rng
 -GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000
 -LOOPEXT=0.000 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000
 -XGAPEXT=0.500 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000
 -YGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1
 -MATRIX=blosum62 -TRANS=human40 cdi -LIST=45 -DOCALLIGN=200
 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
 -OUTFMT=pfs -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
 -USER=US09171553_@GNI_1.526 -NCPU=6 -TCPU=3 -LONGLOG
 -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-171-553b-10
 Query length: 657
 Database: N_Geneseq_1101:*
 Database sequences: 930621
 Database length: 42862619
 Search time (sec): 338.200000

score_list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
/SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:AAV09703		3533.00	4463.44	1.5e-240	1974
/SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:AAV82749		3484.00	4395.62	9.2e-237	3482
/SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:AAV09699		3445.50	4338.21	1.5e-233	8196
/SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:AAV09700		3444.50	4336.93	1.7e-233	8209
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:AAV82748		2631.50	3318.07	9.7e-177	2462
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAV67022		2549.50	3207.15	1.5e-170	4918
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAV67018		2546.50	3212.47	7.4e-171	1980
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAV77726		2537.50	3187.93	1.7e-169	7333
/SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:AAV77483		2537.50	3187.85	1.7e-169	7393
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAV67023		2514.50	3158.05	7.9e-168	7873
/SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:AAV74811		2402.00	3015.16	7.2e-160	8060
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAV77725		2402.00	3015.16	7.2e-160	8060
/SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:AAV74884		2395.00	3006.50	2.2e-159	7892
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAV77727		2395.00	3006.20	2.3e-159	8132
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAV51087		2169.00	2733.68	3.4e-144	2000
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAV51086		1675.00	2110.19	1.8e-109	1493
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAV46343		1558.50	1958.90	4.9e-101	2129
/SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:AAV90691		1544.50	1930.25	1.9e-99	6312
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAV80076		1526.00	1918.04	9.3e-99	2055
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAV46344		1522.50	1913.25	1.7e-98	2129
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAV45540		1522.50	1899.87	9.5e-98	8088
/SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT:AAV07371		1522.50	1899.33	1.0e-97	8535
/SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:AAV66507		1506.50	1893.57	2.1e-97	2004
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAV89080		1488.00	1870.14	4.3e-96	1998
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:AAV69750		1478.00	1843.40	1.3e-94	8655
/SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT:AAV81042		1329.00	1654.21	4.6e-84	8323
/SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:AAV80058		1329.00	1654.21	4.6e-84	8323
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:AAV07063		1329.00	1654.21	4.6e-84	8323
/SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT:AAV58700		1315.00	1636.46	4.5e-83	8323
/SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:AAV30412		1288.50	1616.06	6.1e-82	2229
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:AAV83509		1285.50	1613.36	8.7e-82	1998
/SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:AAV30413		1285.50	1612.06	1.0e-81	2274
/SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:AAV30414		1285.50	1611.72	1.1e-81	2352
/SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:AAV30403		1285.50	1610.55	1.2e-81	2643
/SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:AAV30406		1284.00	1609.95	1.3e-81	2322
/SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:AAV30407		1284.00	1609.76	1.4e-81	2367
/SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:AAV90690		1284.00	1600.33	4.6e-81	6061
/SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT:AAV08452		1284.00	1597.14	6.9e-81	8332
/SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT:AAV07374		1284.00	1597.14	6.9e-81	8332
/SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:AAV05061		1284.00	1597.14	6.9e-81	8332

/SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:AAV42715 + 1284.00 1597.14 6.9e-81
 /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:AAV19551 + 1284.00 1597.14 6.9e-81
 /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:AAV15346 + 1284.00 1597.14 6.9e-81
 /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:AAV28143 + 1284.00 1597.14 6.9e-81
 /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAV40297 + 1284.00 1597.14 6.9e-81

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:AAV09703

seq_documentation_block:

ID AAV09703 standard; DNA; 1974 BP.

XX AAV09703;

DT 20-MAY-1998 (first entry)

DE Porcine retrovirus Raji clone ENV DNA.

XX Porcine retrovirus; PoEV; ENV protein; envelope protein; vaccine;

KW diagnosis; xenotransplantation; prophylactic; therapeutic; ds.

XX Porcine retrovirus.

XX Key Location/Qualifiers

FT CDS 1..1974

FT /tag= a

FT /product= ENV protein

FT /note= "envelope protein"

XX W09740167-AL.

XX 30-OCT-1997.

XX 18-APR-1997; 97WO-GB01087.

XX 10-FEB-1997; 97GB-0002668.

XX 19-APR-1996; 96GB-0008164.

XX (IMUT-) IMUTRAN LTD.

XX (QONE-) Q-ONE BIOTECH LTD.

XX Galbraith DN, Haworth C, Lees GM, Smith KT;

XX WPI; 1997-535851/49.

XX P-PSDB; AAV39274.

XX Polynucleotide encoding porcine retrovirus expression product -

PT useful to develop products for use in vaccines, diagnosis and

PT xeno-transplantation

PS Claim 6; Fig 4; 69pp; English.

XX This sequence encodes the porcine retrovirus envelope (ENV) protein

CC isolated from the human cell line Raji. Such viral proteins can be used

CC to develop viral vaccines, antisense nucleic acids, ribozymes and other

CC antiviral agents. They can also be used in xeno-transplantation

CC technology and as diagnostic tools.

XX Sequence 1974 BP; 534 A; 489 C; 483 G; 468 T; 0 other;

alignment_scores:

Quality: 3533.00 Length: 657
 Ratio: 5.377 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-171-553B-10 x AAV09703

Align seg 1/1 to: AAV09703 from: 1 to: 1974

1 MetHisProThrLeuSerTrpArgHisLeuProThrArgGlyCyluPr 17

|||||

1 ATGCATCCACGTTAAAGCTGGCGCCACCTCCCGACCTCGGGGTGGAGGCC 50

17 oLysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuT 34
|||||
51 GAAAGACATGAGAAATCCCTTAAGCTTGGCTCCCATCGCTGGTTCCTTA 100
34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSer 50
|||||
101 CTCCTAACATACTCCCGAGCCAGTAGTAACGGCTTATAGACAGCTCG 150
51 AsnProHisArgProLeuSerProThrTrpLeuIleIleAspProAspTh 67
|||||
151 AACCCCCATAGACCTTTATCCCTACCTGGCTGATTTATGACCTGTATAC 200
67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrpT 84
|||||
201 GGGTGTCACTGTAAATAGACATCGAGGTGTGCTCCTAGAGCGACCTGGT 250
84 rpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
|||||
251 GGCCCTGAACGTGATTTCTGCCCTCCCATTTGATTAAACCCGCTGTAAAGC 300
101 ThrProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117
|||||
301 ACACCTCCCACTAGTCCGTAGTTATGGTTCTATTGCTGCCAGGCAC 350
117 rGluLysGluLysTyrCysGlySerGlySerGlyCluSerPheCysArgT 134
|||||
351 AGAAGAAAGAAATACTGTGGGGTTCGTGGGGAATCTCTTCTGTAGGAGAT 400
134 rpSerCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150
|||||
401 GGAGCTGCTCACCTCCAAACGATGGAGACTGGAAATGGCCCATCTCTCTC 450
151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrLy 167
|||||
451 CAGGACCGGGTAAATTTCTCTTTGTCAATTCGGCGCGGCAAGTACAA 500
167 sMetMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAspT 184
|||||
501 AATGATGAACATATATAAGATAGAGCTGCTCCCATACAGCTTAGATT 550
184 yrLeuLysIleSerPheThrGluLysGlyLysGlnGluAsnIleGlnLys 200
|||||
551 ATCTAAAGATAAGTTTCACTCAAAAGGAAACAGGAAATATTCAAAG 600
201 TrpIleAsnGlyMetSerTrpGlyIleValPheTyrLysTyrGlyGly 217
|||||
601 TGGATAAATGGTATGAGCTGGGGAATAGTTTTTATAAATATGCGGGG 650
217 yAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGluP 234
|||||
651 AGCAGGTGCCACTTTAACCATTCGCCCTTAGGATAGACGGGGGACAGAC 700
234 roProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProPro 250
|||||
701 CCCCTGTGGCAATGGACCCGATAAAGTACTGGCTGAACAGGGGCCCCCG 750
251 AlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuAr 267
|||||
751 GCCCTGGAGCCACCGCACTAATTCGCCGTGCCCAATTAACCTCGCTGCG 800
267 gProAspIleThrGlnProProSerAsnSerThrThrGlyLeuIleProT 284
|||||
801 GCCTGACATACACAGCGGCTAGCAACAGTACCACCTGGATTGATTCCTA 850
284 hrAsnThrProArgAsnSerProGlyValProValLysThrGlyGlnArg 300
|||||
851 CCAACAGCCCTAGAACTCCCGAGGTCTTCTGTTTATAGACAGGACAGAGA 900
301 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspPr 317
|||||
901 CTCCTTCAGTCTCATCCAGGGAGCTTTCCAAAGCCATCAACTCCAGCCACC 950

317 oAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTyrT 334
|||||
951 TGATGCCACTTCTTCTTGTGGCTTTGTCTATCTCAGGGCCCTCCTTATT 1000
334 yrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArg 350
|||||
1001 ATGAGGGATGGCTAAAGAAAGAAATTCATGTGACCAAGAGCATAGA 1050
351 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSe 367
|||||
1051 AATCAATGTACATGGGGTCCCGAAATAAGCTTACCCTCAGTCAAGTTTC 1100
367 rGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeuC 384
|||||
1101 CGGAGAGGGACATGCTAGTAAAGCTCCCATCCCAACACACCTTT 1150
384 ystYrSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeuVal 400
|||||
1151 GCTATAGTACTGTGTTTATGAGCAGGCCCTCAGAAATCAGTATTAGTA 1200
401 ProGlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysVa 417
|||||
1201 CCTGGTTATAACAGGTGGTGGCATGCAATACTGGGTAAACCCCTGTGT 1250
417 lSerThrSerValPheAsnGlnSerLysAspLeuCysValMetValGlnI 434
|||||
1251 TTCACCTCAGTCTTCAACCAATCCAAGATTGTGTGTCATGGTCCAAA 1300
434 leValProArgValTyrTyrHisProGluGluValValLeuAspGluTyr 450
|||||
1301 TCGTCCCCGAGTGTACTACCATCTCGAGGAGTGTCTTGTATGAATAT 1350
451 AspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAl 467
|||||
1351 GACTATCGGTATAACCGACCAAAAGAGAACCCGTATCCTTACCCTAGC 1400
467 aValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrAlaA 484
|||||
1401 TGTAAATGCTCGGATTAGGACGGCCGTGGCGTAGGAACAGGACAGCTG 1450
484 laLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuGlyGluLeuHis 500
|||||
1451 CCCTGATCACAGGACACAGCAGCTAGAGAAAGGACTTGGTGAGCTACAT 1500
501 AlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnLe 517
|||||
1501 GCGGCCATGACAGAGATCTCCGAGCCTTAAAGAGTCTGTAGCAACCT 1550
517 uGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 534
|||||
1551 AGAAGAGTCCCTGACTTCTTGTCTGAGTGGTTCTTACAGAACCCGAGGG 1600
534 lyLeuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaAlaLeuLys 550
|||||
1601 GATTAGATCTGCTGTTCTTAAGAGAGGTGGGTATGTGACGCCCTAAAA 1650
551 GluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMe 567
|||||
1651 GAAGAATGTTGCTTCTATGTAGTCACTCAGGAGCCATCAGAGACTCCAT 1700
567 tAsnLysLeuArgLysLysLeuArgArgArgGluArgGluAlaA 584
|||||
1701 GAACAAGCTTAGAAAAAAGTTAGAGAGGGCTCGAAGGGAAAGAGAGGCTG 1750
584 spGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThrThr 600
|||||
1751 ACCAGGGTGGTTGAAGGATGGTTCAACAGGCTCCTCTGGATGACCACC 1800
601 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeuTh 617
|||||
1801 CTGCTTCTGCTCTGACGGGGCCCTAGTAGTCTGCTCTCTCTTACTTAC 1850
617 rValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgv 634

```

|||||
1851 AGTTGGGCTTGTCTAAATTAATAGGTTTGTTCCTTTGTTAGAGAACGAG 1900
634 alSerAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeuLeu 650
|||||
1901 TGAGTGCAGTCAGATCATGTACTTAGGCAACAGTAGCAAGGCGCTTCTG 1950
651 SerGlnGlyGluThrAspLeu 657
|||||
1951 AGCCAAGGAGAACTGACCTC 1971

```

seq_name: /SDS2/gcddata/geneseq/geneseqn/NA1999.DAT.AAV82749

seq_documentation_block:

ID AAV82749 standard; DNA; 3482 BP.

AC AAV82749;

XX

DT 25-FEB-1999 (first entry)

XX Pig endogenous retrovirus (PERV)-B envelope (env) gene region.

XX Pig endogenous retrovirus; PERV-A; envelope protein; PERV-B; subtype;
 KW probe; primer; detection; retrovirus; human tissue; xenotransplant;
 KW primary porcine tissue; human cell line; porcine cell line; ss.
 XX

OS Pig endogenous retrovirus.

XX Key Location/Qualifiers

PH CDS 911...2884

FT /*tag= a

FT /product= envelope_protein

XX W09853104-A2.

XX 26-NOV-1998.

XX 18-MAY-1998; 98WO-GB01428.

XX 16-MAY-1997; 97GB-0010154.

XX (MEDI-) MEDICAL RES COUNCIL.

XX Stoye JP, Weiss RA;

XX WPI; 1999-045324/04.

XX P-PSDB; AAW85452.

XX Newly isolated nucleic acid probe capable of hybridising to either
 PT the PERV-A or PERV-B env gene - useful in the detection of
 PT retroviruses, and their subtypes, in a sample of porcine/human
 PT tissue

XX Claim 3; Page 21-23; 36pp; English.

XX The present sequence encodes a pig endogenous retrovirus (PERV)-B
 CC envelope protein. PERV exists in two different subtypes, PERV-A and
 CC PERV-B. The differences are reflected in sequence divergence in the
 CC envelope genes. Probes and primers can be derived from the envelope
 CC (env) genes of PERV-A and PERV-B. The probes and primers are used in
 CC a method to detect retroviruses in a sample of porcine/human tissue,
 CC particularly primary porcine tissue and human cell lines that have been
 CC cultivated in the presence of a porcine cell line, or human tissue from
 CC a patient with a xenotransplant. Subtype of PERV in a sample containing
 CC one of the PERV env genes can also be determined.

XX Sequence 3482 BP; 927 A; 854 C; 867 G; 834 T; 0 other;

alignment_scores:

Quality: 3484.00 Length: 657
 Ratio: 5.319 Gaps: 0
 Percent Similarity: 99.696 Percent Identity: 98.478

```

alignment_block:
US-09-171-553B-10 x AAV82749
Align seg 1/1 to: AAV82749 from: 1 to: 3482

1 MetHisProThrLeuSerTrpArgHisLeuProThrArgGlyGlyGluPr 17
|||||
911 ATGCATCCACAGTTAAGCTGGCCACACTCCCGACTCGGGGTGGAGAGCC 960

17 oLysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuT 34
|||||
961 GAAAGAGACTGAGAAATCCCTTAAGCTTGGCTCCATCGCTGCTTCTTAA 1010

34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuLeuLeuSerSer 50
|||||
1011 CTCTAACAAATACTCCCGAGCCAGTAGTAAGCGCTTATAGACAGCTCG 1060

51 AsnProHisArgProLeuSerProThrTrpLeuIleIleAspProAspTh 67
|||||
1061 AACCCCATAGACCTTTATCCCTTACCTGGCTGATTATTGACCTGTATAC 1110

67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrpT 84
|||||
1111 GGGTGTCACTGTAAATAGCACTCGAGGTGTGCTCTAGAGGCACCTGGT 1160

84 rpProGluLeuHisPheCysLeuArgLeuLeuLeuAsnProAlaValLys 100
|||||
1161 GGCCTGAATGCAATTTCTGCTCCGATTGATTAAACCCGCTGTAAAGAC 1210

101 ThrProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117
|||||
1211 ACACCTCCCAACCTAGTCGTAGTATGGGTCTATTGCTGCCAGGCAC 1260

117 rGluLysGluLysTyrCysGlySerGlyGluSerPheCysArgArgT 134
|||||
1261 AGAGAAAGAGAAATACTGTGGGGTCTGGGAAATCCTTCTGTAGGAGAT 1310

134 rpSerCysValThrSerAsnAspGlyAspTrpLysTrpProLysSerLeu 150
|||||
1311 GGAGCTGCGTCACTCCCAACGATGGAGACTGGAAATGGCCGATCTCTCTC 1360

151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrI 167
|||||
1361 CAGGACCGGGTAAATTTCTTCTCAATTCCGGCCCGGGCAAGTACAA 1410

167 sMetMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAspT 184
|||||
1411 AGTGATGAAACTATATAAAGATAAGAGCTGCTCCCATCAGACTTAGATT 1460

184 yrLeuLysIleSerPheThrGluLysGlyLysGlnGluAsnIleGlnLys 200
|||||
1461 ATCTAAAGATAAGTTTCACTGAAAGAGGAAACAGGAAATAATTCAAAG 1510

201 TrpIleAsnGlyMetSerTrpGlyIleValPheTyrLysTyrGlyGly 217
|||||
1511 TGGATAAATGGTATGAGCTGGGAATAGTTTTTATAATATGGCGGGG 1560

217 yAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGluP 234
|||||
1561 AGCAGGGTCCACTTTAAACATTTCGCTTAGGATAGAGACGGGACAGAAC 1610

234 roProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProPro 250
|||||
1611 CCCCTGTGCACTGGGACCCGATAAGTACTGGCTCAACAGGCGGCCCG 1660

251 AlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuAr 267
|||||
1661 GCCCTGGAGCCACCGCATTAACCTTGGCGTGCCCAATTAACCTCGCTCG 1710

267 gProAspIleThrGlnProProSerAsnSerThrThrGlyLeuLeuProT 284
|||||
1711 GCCTGACATTAACACAGCCGCTTAGCAACGGTACCACCTGGATTGATTCT 1760

```

```

284 hrAsnThrProArgAsnSerProGlyValProValLysThrGlyClnArg 300
|||||
1761 CCAACACGCCATAGAAACTCCCGAGGTCTCTGTAAAGACAGGACAGAGA 1810
|||||
301 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspPr 317
|||||
1811 CTCCTTCAGTCTCATCCAGGGAGCTTCCACGCCATCAACTCCACGACCC 1860
|||||
317 oAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTyr 334
|||||
1861 TCATGCCACTCTCTCTGTGTGCTTGTCTATCTCCACGGGCTCTCTATT 1910
|||||
334 yGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArg 350
|||||
1911 ATGAGGGGATGGCTAAAGAAAGAAATTCATGTGACCAAGACATAGA 1960
|||||
351 AsnGlnCysThrTrpGlySerArgAsnLysLysLeuThrGluValSe 367
|||||
1961 AATCAATGTACATGGGGTCCGGAATTAAGCTTACCCCTCACTGAAGTTTC 2010
|||||
367 rGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeuC 384
|||||
2011 CGGGAAGGGACATCATAGGAAGAGCTCCCCATCCCAACCAACCTTT 2060
|||||
384 ysTyrSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeuVal 400
|||||
2061 GCTATAGTACTGTGTTTATGAGCAGGCTCAGAAATCAGTATTAGTA 2110
|||||
401 ProGlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysVa 417
|||||
2111 CTGGTTTAAACAGTGGTGGCATGCAATCTGGTTAAACCCCTGTGT 2160
|||||
417 lSerThrSerValPheAsnGlnSerLysAspLeuCysValMetValGlnI 434
|||||
2161 TTCCACCTCAGTCTCAACCAATCCAAAGATTCTGTGTCATGTGCCAAA 2210
|||||
434 leValProArgValTyrTrpHisProGluGluValValLeuAspGluTyr 450
|||||
2211 TGTCTCCCGAGTGTACTACCATCTCGAGAGTGGTCTCTGATGAATAT 2260
|||||
451 AspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAl 467
|||||
2261 GACTATCGGTATACCGACCAAAAGAGAACCCGTATCCCTTACCCTAGC 2310
|||||
467 aValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAlaA 484
|||||
2311 TGTAAATGCTCGGATTAGGACGGCGTGTGGCTAGGAACAGGACAGCTG 2360
|||||
484 laLeuIleThrGlyProGlnGlnLeuLysGlyLeuGlyGluLeuHis 500
|||||
2361 CCTGATCACAGGACACAGCAGCTAGAGAAAGGACTTGGTGAAGCTACAT 2410
|||||
501 AlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnLe 517
|||||
2411 GCGGCCATGACAGACATCTCGAGCCCTAGAGAGTCTGTATCAACCT 2460
|||||
517 uGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 534
|||||
2461 AGAAGAGTCCCTGACTCTTTGCTGAAAGTGGTCTACAGAACCGGAGGG 2510
|||||
534 lYLeuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaAlaLeuLys 550
|||||
2511 GATTAGATCTGCTGTTTCTAAGAGAAAGTGGGTATGTCAGACCTTAAA 2560
|||||
551 GluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMe 567
|||||
2561 GAAGAATGTGCTCTCTATGTAGATCACTCAGAGGCCATCAGACTCCAT 2610
|||||
567 tAsnLysLeuArgLysLysLeuGluArgArgArgGluAlaA 584
|||||
2611 GAGCAAGCTTAGAGAAAGGTTAGAGAGCGCTCGAAGGGAAGAGAGCTG 2660
|||||

```

```

584 spGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThrThr 600
|||||
2661 ACCAGGGTGGTTTGAAGATGGTTCAACAGGTCTCTTGGATGACCACC 2710
|||||
601 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeu 617
|||||
2711 CTGCTTCTCTCTCTGACGGGACCCCTAGTAGTCTCTCTCTTACTTAC 2760
|||||
617 rValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgV 634
|||||
2761 AGTGGGCCCTTGCTTAATTAATAGTTTGTGCTTTGTAGAGAACGAG 2810
|||||
634 alSerAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeuLeu 650
|||||
2811 TGAGTGCGAGTCCAGATCATGTACTTAGGCACAGTACCAAGGCCTTCTG 2860
|||||
651 SerGlnGlyGluThrAspLeu 657
|||||
2861 AGCCAAGGAGAAACTGACCTC 2881
|||||
seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:AAV09699
seq_documentation_block:
ID AAV09699 standard; DNA; 8196 BP.
XX
AC AAV09699;
XX
DT 19-MAY-1998 (first entry)
XX
DE Porcine retrovirus DNA encoding, GAG, POL and ENV.
XX
KW Porcine retrovirus; PoEV; POL protein; ENV protein; GAG protein;
XX vaccine; diagnosis; xenotransplantation; prophylactic; therapeutic; ds.
XX
OS Porcine retrovirus.
XX
FH Key Location/Qualifiers
CDS 576..2126
FT /*tag= a
FT /product= GAG protein
FT /note= "viral core polypeptide"
CDS 2143..5733
FT /*tag= b
FT /product= POL protein
FT /note= "polymerase peptide sequence as given in
FT 5606..7576
FT /*tag= c
FT /product= ENV protein
FT /note= "envelope protein"
XX
PN WO9740167-A1.
XX
PD 30-OCT-1997.
XX
PF 18-APR-1997; 97WO-GB01087.
XX
PR 10-FEB-1997; 97GB-0002668.
PR 19-APR-1996; 96GB-0008164.
XX
PA (IMUT-) IMUTRAN LTD.
PA (QONE-) Q-ONE BIOTECH LTD.
XX
PI Galbraith DN, Haworth C, Lees GM, Smith KT;
XX
WPI; 1997-535851/49.
XX
DR Polynucleotide encoding porcine retrovirus expression product -
XX useful to develop products for use in vaccines, diagnosis and
XX xeno-transplantation
XX
PS Claim 4; Fig 2; 69pp; English.
XX

```

CC This DNA sequence encodes the porcine retrovirus (PoEV) virion core
 CC polypeptide (GAG), polymerase (POL), and envelope (ENV) proteins. These
 CC proteins can be used to develop viral vaccines, antisense nucleic acids,
 CC ribozymes and other antiviral agents. They can also be used in
 CC xeno-transplantation technology and as diagnostic tools.
 XX
 SQ Sequence 8196 BP; 2165 A; 2061 C; 2147 G; 1820 T; 3 other;

alignment_scores:

Quality: 3445.50 Length: 658
 Ratio: 5.276 Gaps: 2
 Percent Similarity: 99.240 Percent Identity: 98.936

alignment_block:

US-09-171-553B-10 x AAV09699 ..

Align seg 1/1 to: AAV09699 from: 1 to: 8196

1 MethisProThrLeuSerTrpArgHisLeuProThrArgGlyGlyGluPr 17
 |||||
 5606 ATGCATCCACGTTAAGCYGGGCCACCTCCGACTCGGGTGGAGAGCC 5655
 17 OlysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuT 34
 |||||
 5656 GAAAGACCTGAGAAATCCCTTAAGCTTCGCTCCATCGCTGGTTCCTTA 5705
 34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50
 |||||
 5706 CTCTACATAACTCCCGAGCGCAGTAGTAAGCGCTTATAGACAGCTCG 5755
 51 AsnProHisArgProLeuSerProThrTrpLeuIleAspProAspTh 67
 |||||
 5756 AACCCCATAGACCTTTATCCCTTACCTGGCTGATTATTGACCTGATAC 5805
 67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrpT 84
 |||||
 5806 GGGTGTACATGTAATAGACACTCGAGGTGTCTCTCTAGAGCACCTGTG 5855
 84 rpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
 |||||
 5856 GGCTGAACCTGATTTCTCCCTCCGATGATTAAACCCGCTGTTAARGAG 5905
 101 ThrProProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117
 |||||
 5906 ACACCTCCCAACCTAGTCCGTAGTTATGGTCTATTGGTGTCCCGAGCAC 5955
 117 rGluLysGluLysTyrCysGlySerGlySerGlyGluSerPheCysArgT 134
 |||||
 5956 AGAGAAAGAGAAATACTGTGGGGGTCTGGGGAATCCTTCTGTAGGAGAT 6005
 134 rpSerCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150
 |||||
 6006 GGAGCTGCCTACCTCCACGATGAGACTGGAATGGCCGATCTCTCTC 6055
 151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrLy 167
 |||||
 6056 CAGGACGGGTAAATTCCTTTGTCAATTCGCGCCGGGCAAGTACAA 6105
 167 sMetMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAspT 184
 |||||
 6106 AATGATGAACATATATAAGATAAGAGCTGTCCCATCAGACTTAGATT 6155
 184 yrLeuLysIleSerPheThrGluLys.GlyLysGlnGluAsnIleGlnLy 200
 |||||
 6156 ATCTAAAGATAAGTTTCACT...GAAGAGAAACAGGAAATATTCAAAA 6202
 200 sTrpIleAsnGlyMetSerTrpGlyIleValPheTyrLysTyrGlyGlyG 217
 |||||
 6203 GTGGATAAATGATAGCTGGGGAATAGTTTATTATTA.TATGGCGGGG 6251
 217 lyAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGlu 233
 |||||

6252 GAGCAGGGTCCACTTTAAACCATTCGCCTTAGGATAGACACGGGACAGAA 6301
 234 ProProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProPr 250
 |||||
 6302 CCCCTGTGGCAATGGGACCCGATAAAGTACTGGCTGAACAGGGGCCCCC 6351
 250 oAlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuA 267
 |||||
 6352 GGCCCTGGAGCCACCGCATAACTTGCCTGGTCCCAATTAACTCGCTGC 6401
 267 rgProAspIleThrGlnProProSerAsnSerThrThrGlyLeuIlePro 283
 |||||
 6402 GGCCTGACATAACACAGCGGCTAGCAACAGTACCACCTGGATTGATTCT 6451
 284 ThrAsnThrProArgAsnSerProGlyValProValLysThrGlyGlnAr 300
 |||||
 6452 ACCAACACGCTTAGAACTCCCGAGGTGTCTGTAAAGACAGGACAGAG 6501
 300 gleuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspP 317
 |||||
 6502 ACTCTTCAGTCTCATCCAGGGAGCTTTCACAGCCATCAACTCCACGACC 6551
 317 roAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTyr 333
 |||||
 6552 CTGATGCCACTTCTTCTTGGTCTTGTCTATCTCAGGGCTCTCTTAT 6601
 334 TyrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisAr 350
 |||||
 6602 TATGAGGGGATGGCTAAAGAAAGAAATTCATGTGACCAAGAGACATAG 6651
 350 gAsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValS 367
 |||||
 6652 AAATCAATGTACATGGGGGTCCCGAATAAGCTTACCTCTCACTGAAGTT 6701
 367 erGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeu 383
 |||||
 6702 CCGGAGGGGACATGCATAGAAAAGCTCCCCATCCCAACACACTT 6751
 384 CysTrpSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeuVa 400
 |||||
 6752 TGCTATAGTACTGTGTTTATGACGAGGCTCAGAAAATCAGTATTATGT 6801
 400 lProGlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysV 417
 |||||
 6802 ACCTGGTTATAACAGGTGGTGGGCATCAATACTGGGTAAACCCCTGTG 6851
 417 alSerThrSerValPheAsnGlnSerLysAspLeuCysValMetValGln 433
 |||||
 6852 TTTCCACCTCAGTCTTCAACCAATCCAAAGATTTCTGTGTCTCATGGTCAA 6901
 434 lIleValProArgValTyrTyrHisProGluGluValValLeuAspGluTy 450
 |||||
 6902 ATCGTCCCCCGAGTGTACTACCATCTCGAGGAGGTGCTTGTATGAATA 6951
 450 rAspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuA 467
 |||||
 6952 TGACTATCGGTATAACGACCAACCAAGAGAACCGGTATCCCTTACCCCTAG 7001
 467 laValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAla 483
 |||||
 7002 CTGTAATGCTCGGATTAGGACGCGCTTGGCGTAGGAACAGGACAGCT 7051
 484 AlaLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuGlyGluLeuHi 500
 |||||
 7052 GCCCTGATACAGGACCACAGCAGCTAGAGAAAGGACTTGTGTGAGCTACA 7101
 500 sAlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnL 517
 |||||
 7102 TGGCGCCATGACAGACATCTCCGAGGCTTAAAGAGAGTCTGTAGCAACC 7151
 517 euGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 533
 |||||
 7152 TAGAAGAGTCCCTGACTTCTTTGTCTGAGTGTCTTCTACAGAACCGGAGG 7201

XX

PI Galbraith DN, Haworth C, Lees GM, Smith KT;

XX

WPI: 1997-535851/49.

DR

P-PSDB; AAW39271; AAW39272; AAW39273.

XX

PT Polynucleotide encoding porcine retrovirus expression product -
 PT useful to develop products for use in vaccines, diagnosis and
 PT xeno-transplantation

XX

Claim 4; Fig 3; 69pp; English.

XX

CC This DNA sequence encodes the porcine retrovirus (PoEV) virion core
 CC polypeptide (GAG), polymerase (POL) and envelope (ENV) proteins and
 CC also includes the Long Terminal Repeat (LTR). These proteins can be used
 CC to develop viral vaccines, antisense nucleic acids, ribozymes and other
 CC antiviral agents. They can also be used in xeno-transplantation
 CC technology and as diagnostic tools.

XX

Sequence 8209 BP; 2168 A; 2064 C; 2154 G; 1823 T; 0 other;

alignment_scores:

Quality: 3444.50 Length: 658

Ratio: 5.275 Gaps: 2

Percent Similarity: 99.240 Percent Identity: 98.936

alignment_block:

US-09-171-553B-10 x AAV09700 ..

Align seg 1/1 to: AAV09700 from: 1 to: 8209

1 MethisProThrLeuSerTrpArgHisLeuProThrArgGlyGlyGluPr 17

5620 ATGCATCCACGTTAAGCGCGCCACCTCCCGACTCGGGGTGGAGACC 5669

17 OlyArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuT 34

5670 GAAAGAGCTGAGAAATCCCTTAAAGCTTCGGCTCCATCGCTGTTCCCTTA 5719

34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50

5720 CTCCTAACAAATACCTCCCGAGCCAGTAGTAAACGCCCTTATAGACAGCTCG 5769

51 AsnProHisArgProLeuSerProThrTrpLeuIleIleAspProAspTh 67

5770 AACCCCATAGACCTTTATCCCTTACCTGGCTGATTTATGACCTGATAC 5819

67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrpT 84

5820 GGGTGTCACTGTAATAGCACTCGAGGTGTGCTCCCTAGAGGCACCTGGT 5869

84 rpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100

5870 GGCCCTGAACCTGCAATTCCTGCTCCGATTGATTAAACCCCGCTGTAAAGC 5919

101 ThrProProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117

5920 ACACCTCCCAACCTAGTCCGTAGTTATGGGTTCTATTGCTGCCAGGCAC 5969

117 rGluLysGluLysTyrCysGlyGlySerGlyGluSerPheCysArgArgT 134

5970 AGAGAAAGAGAAATACCTGTGGGGTTCTGGGGAATCCTTCTGTAGGAGAT 6019

134 rpSerCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150

6020 GGAGCTCGGTCACTCCCAAGATGGAGACTGGAAATGGCGGATCTCTCTC 6069

151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrLy 167

6070 CAGGACCGGTAAATTCCTCTTTGTCATTCCTCGGCCCGGCAAGTACAA 6119

167 sMetMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAspT 184

6120 AATGATGAACATATATAAGATAAGAGCTGCTCCCATCATGACTTACATT 6169

184 YrLeuLysIleSerPheThrGluLys_GlyLysGlnGluAsnIleGlnLy 200

6170 ATCTAAAGATAAGTTTCACT...GAAAGGAAACAGAGAAATATCAAAA 6216

200 sTrpIleAsnGlyMetSerTrpGlyIleValPheTyrLysTyrGlyGlyG 217

6217 GTGGATAAATGGTATGAGCTGGGGAATAGTTTTTTATTA.TATGGCGGG 6265

217 lyAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGlu 233

6266 GAGCAGGGTCCACTTTAACCATTCGCTTAGGATAGAGAGGGGACAGAA 6315

234 ProProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProPr 250

6316 CCCCCTGTGGCAATGGGACCCGATAAAGTACTGGCTGAACAGGGGCCCC 6365

250 oAlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuA 267

6366 GGCCTGGAGCCACCGCATAACTTGGGFGCCCCCAATTAACTCTCGTCG 6415

267 rgProAspIleThrGlnProProSerAsnSerThrThrGlyLeuIlePro 283

6416 GGCCTGACATAACACAGCGCCCTAGCAACAGTACCACCTGGATTGATCT 6465

284 ThrAsnThrProArgAsnSerProGlyValProValLysThrGlyGlnAr 300

6466 ACCAACCGCTAGAAACTCCCGAGGTGTCTGTTTAAAGACAGGACAGAG 6515

300 gLeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAsp 317

6516 ACTCTTCAGTCTCATCCAGGAGCTTCCCAAGCCATCAACTCCACGACC 6565

317 roAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTyr 333

6566 CTGATGCCACTTCTTCTGTGGCTTTGTCTATCTCTCAGGGCTCTCTTAT 6615

334 TyrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisAr 350

6616 TATGAGGGGATGGCTAAAGAAAGAAATTCATGTGACCAAGAGAGCATAG 6665

350 gAsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluVal 367

6666 AATCAATGTACATGGGGTCCGGAATAAGCTTACCTCCTCAGTGAATTT 6715

367 erGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeu 383

6716 CCGGGAAGGGGACATGCATAGGAAAGCTCCCCCATCCCAACCAACACCTT 6765

384 CysTyrSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeuVa 400

6766 TGCTATAGTACTCTGCTTTATGAGCAGGCTCAGAAATTCAGTATTTAGT 6815

400 lProGlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysV 417

6816 ACCTGGTTATAACAGGTGGGCGATGCANTACTGGGTTAACCCCTGTG 6865

417 alserThrSerValPheAsnGlnSerLysAspLeuCysValMetValGln 433

6866 TTTCCACCTCAGTCTTCAACCAATCCAAAGATTTCTGTGTGTCATGTC 6915

434 IleValProArgValTyrTyrHisProGluGluValValLeuAspGluTy 450

6916 ATCGTCCCCCGAGTGTACTTACCATCTCGAGGAAGTGGTCTTGTATGA 6965

450 rAspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuA 467

6966 TCACATATCGGTATACCCGACCAAGAGAACCCGTCATCCCTTACCCCT 7015

467 laValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAla 483

7016 CTGTATAGTCGATTAGGACGGCGTTGGCTAGGAACAGGGACAGCT 7065
 484 AlaLeuIleThrGlyProGlnLeuGluGlyLeuGlyLeuGluHi 500
 |||||
 7066 GCCTGATCACAGGACACAGCAGCTAGAGAAGGACTTGGTGAGCTACA 7115
 500 sAlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnL 517
 |||||
 7116 TGGGGCCATCAGACAGATCTCGAGCCTTAAAGGAGTCTGTTAGCAACC 7165
 517 euGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 533
 |||||
 7166 TAGAAGAGTCCCTGACTCTTGTCTGAAGTGGTTCTACAGAACCGGAGG 7215
 534 GlyLeuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaAlaLeuLy 550
 |||||
 7216 GGATTAGATCTGCTGTTCTTAAGAGAAGGTGGTTATGTGACGCTTAA 7265
 550 sGluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerM 567
 |||||
 7266 AGAAGAATGTTGCTTCTATGTATGATCAGCTCAGAGGCCATCAGAGACTCCA 7315
 567 eAsnLysLeuArgLysLysLeuGluArgArgArgArgGluAla 583
 |||||
 7316 TGAACAAGCTTACAAAAAGTTAGAGAGCGCTCGAAGGGAAGAGAGGCT 7365
 584 AspGlnGlyTyrPheGluGlyTrpPheAsnArgSerProTirpMetThrTh 600
 |||||
 7366 GACCAGGGTGGTTTCAAGGATGGTTCAACAGGTCTCCTTGGATGACCAC 7415
 600 rLeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeu 617
 |||||
 7416 CTTGCTTTCTGCTCTGACGGGCGCCCTAGTAGCTCTGCTGCTGTTACTTA 7465
 617 hrValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArg 633
 |||||
 7466 CAGTGGGCTTGCCTTAATTAATAGTTGTTGCTTGTAGAGAACA 7515
 634 ValSerAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeuLe 650
 |||||
 7516 GTGAGTGCAGTCCAGATCATGTTAGTCTAGGCAACAGTACCAAGGCTTCT 7565
 650 uSerGlnGlyGluThrAspLeu 657
 |||||
 7566 GAGCCAAGGAGAACTGACCTC 7587

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:AAV82748

seq_documentation_block:

ID AAV82748 standard; DNA; 2462 BP.

XX AC AAV82748;

XX DT 25-FEB-1999 (first entry)

XX DE Pig endogenous retrovirus (PERV)-A envelope (env) gene region.

XX KW Pig endogenous retrovirus; PERV-A; envelope protein; PERV-B; subtype:
 probe; primer; detection; retrovirus; human tissue; xenotransplant;
 primary porcine tissue; human cell line; porcine cell line; ss.

XX OS Pig endogenous retrovirus.

XX FH Key Location/Qualifiers

FT CDS 211..2193

FT /*tag= a

FT /product= envelope_protein

XX PN W09853104-A2.

XX PD 26-NOV-1998.

XX PF 18-MAY-1998; 98WO-GB01428.

XX 16-MAY-1997; 97GB-0010154.
 XX (MEDI-) MEDICAL RES COUNCIL.
 XX Stoye JP, Weiss RA;
 XX WPI; 1999-045324/04.
 DR P-PSDB; AAW85453.
 XX Newly isolated nucleic acid probe capable of hybridising to either
 PT the PERV-A or PERV-B env gene - useful in the detection of
 PT retroviruses, and their subtypes, in a sample of porcine/human
 PT tissue
 XX Claim 6; Page 20-21; 36pp; English.
 XX The present sequence encodes a Pig endogenous retrovirus (PERV)-A
 CC envelope protein. PERV exists in two different subtypes, PERV-A and
 CC PERV-B. The differences are reflected in sequence divergence in the
 CC envelope genes. Probes and primers can be derived from the envelope
 CC (env) genes of PERV-A and PERV-B. The probes and primers are used in
 CC a method to detect retroviruses in a sample of porcine/human tissue,
 CC particularly primary porcine tissue and human cell lines that have been
 CC cultivated in the presence of a porcine cell line, or human tissue from
 CC a patient with a xenotransplant. Subtype of PERV in a sample containing
 CC one of the PERV env genes can also be determined.
 XX Sequence 2462 BP; 715 A; 573 C; 575 G; 599 T; 0 other;

alignment_scores:

Quality: 2631.50 Length: 679

Ratio: 4.468 Gaps: 9

Percent Similarity: 86.745 Percent Identity: 73.049

alignment_block:

US-09-171-553B-10 x AAV82748 ..

Align seg 1/1 to: AAV82748 from: 1 to: 2462

1 MethisProThrLeuSerTrpArgHisLeuProThrArgGlyGlyGluPr 17
 |||||
 211 ATGCATCCACGTTAAGCGCGCCACCTCCCGATTCCGGGTGGAAGCC 260
 17 oLysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeu 34
 |||||
 261 GAAAGACTGAAATCCCTTAAGCTTCGCTCCATCGCGTGTCTCTTA 310
 34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50
 |||||
 311 CTGTGCAATTAACCTCTCAAGTTAATGGTAAACGCTTGTGGACAGCCCG 360
 51 AsnProHisArgProLeuSerProThrTrpLeuIleAspProAspTh 67
 |||||
 361 AACTCCCATATAACCCCTTATCTCTCACCTGGTTACTTACTGACTCCGGTAC 410
 67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84
 |||||
 411 AGGTATTAAATTAACAGCACTCAAGGGGAGGCTCCCTTGGGACCTGGT 460
 84 rpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaVal..... 98
 |||||
 461 GGCTGTAATTAATATGTCTGCTTCGATCAGTAATCCCTGGTCTCATATGAC 510
 99 LysSerThrProAsnLeuValArgSerTyrGlyPheTyrCysCysPr 115
 |||||
 511 CAGGCCACACCCCGATGACTCCGTGCTTACGGGTTTACGGTTTGCCTC 560
 115 oGly...ThrGluLysGluLysTyrCysGlyGlySerGlyGluSerPheC 131
 |||||
 561 AGGACCCCAATAATGAAGAAATATTGTGGAATATCCCTCAGGATTTCTTT 610

```

131 ysArgArgTrpSerCysValThrSerAsnAspGlyAspTrpLysTrpPro 147
|||||
611 GCAAGCAATGGAGCTGCATAACTTCTAATGATGGAAATGGAAATGGCCA 660
|||||
148 lIeSerLeuGlnAspArgValLysPheSerPheValAsnSer..... 161
|||||
661 GTCTCTCAGCAAGACAGAGTAAGTACTCTTTGTTTAAACAATCTACCAG 710
|||||
162 .....GlyProGlyLysTyrLys.....MetMetL 170
|||||
711 TTATAATCAATTTAATTTATGCGCATGGAGATGGAAAGATGGCAACAGC 760
|||||
170 ysLeuTyrLysAspLys.....SerCysSerProSerAsp 181
|||||
761 GGGTACAAAAGAGATGTACGAAATAAGCAATAAGCTGCTCATCTCGTTAGAC 810
|||||
182 LeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGluAsnII 198
|||||
811 CTAGATTACTTAAATAAAGTTTCACTGAAAAAGGAAACAAGAAATAT 860
|||||
198 eGlnLysTrpIleAsnGlyMetSerTrpGlyIleValPheTyrLysTyrG 215
|||||
861 TCAAAAGTGGTAATAGTATATCTTGGGAATAGTGTAC.....TATG 904
|||||
215 lyGlyGlyAla.....GlySerThrLeuThrIleArgLeuArgIle 228
|||||
905 GAGGCTCTGGGAGAAAGAAAGGATCTCTTCTGACTATTCGCCCTCAGAA 954
|||||
229 GluThrGlyThrGluProProValAlaMetGlyProAspLysValLeuAl 245
|||||
955 GAAACTCAGATGGAACCTCCGGTGTCTATAGGACCAATAAAGGTTTGGC 1004
|||||
245 aGluGlnGlyProProAlaLeuGluProProHisAsnLeuProValProG 262
|||||
1005 CGAACAGGACCTCCAAATCCAAAGAACAG..... 1032
|||||
262 lnLeuThrSerLeuArgProAspIleThrGlnProProSerAsnSerThr 278
|||||
1033 .....AGGCCATCTCCTAACCCCTCTGATTACATACACACC 1068
|||||
279 ThrGlyLeuIleProThrAsnThrProArgAsnSerProGlyValProVa 295
|||||
1069 TCTGGATCAGTCCCACT.....GAGCCTAACATCAGTAT 1103
|||||
295 lLysThrGlyGlnArgLeuPheSerLeuIleGlnGlyAlaPheGlnAlaI 312
|||||
1104 TAAACAGGGGGCAAGCTTTTATGCCCTATCCAGGGAGCTTTCAAGCTC 1153
|||||
312 leAsnSerThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSer 328
|||||
1154 TTAACCTCAGGACCTCCAGAGGCTACCTCTTCTTGGCTTGGCTTAGCT 1203
|||||
329 SerGlyProProTyrTyrGluGlyMetAlaLysGluArgLysPheAsnVa 345
|||||
1204 TCGGGCCCACTTACTACTATGAGGAATGGCTAGAGAGGGAATTCATGT 1253
|||||
345 lThrLysLysHisArgAsnGlnCysThrTrpGlySerArgAsnLysLeuT 362
|||||
1254 GACAAAGGACATAGACCAATGTATGCGGATCCCAAAATAAGCTTA 1303
|||||
362 hrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysAlaProPro 378
|||||
1304 CCCTTACTAGGTTTCTGGAAAGGACCTGCGATAGGATGGTTCGCCCA 1353
|||||
379 SerHisGlnHisLeuCysTyrSerThrValValTyrGluGlnAlaSerGI 395
|||||
1354 TCCCAACCAACACCTTTCTAACACACACTGAAGCCTTTAATCGAACCTCTGA 1403
|||||
395 uAsnGlnTyrLeuValProGlyTyrAsnArgTrpTrpAlaCysAsnThrG 412
|||||
1404 GAGTCAATATCTGTACTCTGTATGACAGAGGTGGGCGATGTAATACTG 1453
|||||
412 lyLeuThrProCysValSerThrSerValPheAsnGlnSerLysAspLeu 428

```

```

1454 GATTAACCCCTTGTGTTCACCTTGTTCACCCAACTAAAGACTTT 1503
|||||
429 CysValMetValGlnIleValProArgValTyrHisProGluGluVa 445
|||||
1504 TCGGTATGTGTCCAAATGTCCCGGGTGTACTACTATCCGCAAAAGC 1553
|||||
445 ValLeuAspGluTyrAspTyrArgTyrAsnArgProLysArgGluProV 462
|||||
1554 AGTCTTGATGATATGACTATAGATATATCGCCAAAGAGAGACCCA 1603
|||||
462 alSerLeuThrLeuAlaValMetLeuGlyLeuGlyThrAlaValGlyVal 478
|||||
1604 TATCCCTGACACTAGCTGTAATGCTCGGATTTGGAGTGGCTCAGCGGTG 1653
|||||
479 GlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeuGluLysGI 495
|||||
1654 GGAACAGGAACGGCTGCCCTAATACAGGACCCCAACAGCTGGAAAGG 1703
|||||
495 yLeuGlyGluLeuHisAlaAlaMetThrGluAspLeuArgAlaLeuLysG 512
|||||
1704 ACTTAGTAACCTACATCGAATTGTACGGAAGATCTCCAGGCCCTAGAAA 1753
|||||
512 luSerValSerAsnLeuGluGluSerLeuThrSerLeuSerGluValVal 528
|||||
1754 AATCTGTCACTAACCTGGAGGAATCCCTAACCTCTTATCTCAAGTGGTT 1803
|||||
529 LeuGlnAsnArgArgGlyLeuAspLeuLeuPheLeuArgGluGlyGlyLe 545
|||||
1804 CTACAGAACAGAGGGGGTTAGATCTGTTATTTCTAAAAGAGAGGGTT 1853
|||||
545 uCysAlaAlaLeuLysGluGluCysCysPheTyrValAspHisSerGlyA 562
|||||
1854 ATGTGTAGCCTTAAAGAGGAATGCTGCTTCTATGTAGTACCTCAGGAG 1903
|||||
562 laIleArgAspSerMetAsnLysLeuArgLysLysLeuGluArgArgArg 578
|||||
1904 CCATCAGAGACTCCATGACCAAGCTTAGAGAAAGGTTAGAGAGGCTCGA 1953
|||||
579 ArgGluArgGluAlaAspGlnGlyTrpPheGluGlyTrpPheAsnArgSe 595
|||||
1954 AGGGAAAGAGAGAGCTGACAGGGGGTGTGAAAGGATGGTTTCAACAGGTC 2003
|||||
595 rProTrpMetThrThrLeuLeuSerAlaLeuThrGlyProLeuValValL 612
|||||
2004 TCCTTGGATGACCACTCTTCTGCTCTGACGGGGCCCCCTAGTAGTCC 2053
|||||
612 euLeuLeuLeuThrValGlyProCysLeuIleAsnArgPheValAla 628
|||||
2054 TGCTCTCTGTTACTTACAGTTGGGCTTGCTTAATTAATAGGTTTGTGGC 2103
|||||
629 PheValArgGluArgValSerAlaValGlnIleMetValLeuArgGlnGI 645
|||||
2104 TTTGTTAGAGAACGAGTAGTGCAGTCCAGATCATGGTACTTAGCAACA 2153
|||||
645 nTyrGlnGlyLeuLeuSerGlnGlyGluThrAspLeu 657
|||||
2154 GTACCAAGGCCCTTCTGAGCCCAAGGAGAAACTGACCTC 2190
|||||

```

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT: AAC67022

seq_documentation_block:

ID AAC67022 standard; DNA; 4918 BP.

XX AC AAC67022;

XX DT 27-MAR-2001 (first entry)

XX DE PERV env protein coding sequence SEQ ID NO: 22.

XX KW Xenotransplantation; infectious agent; vaccine; ds.

XX OS Porcine endogenous retrovirus.

XX WO200071726-A1.
 PN 30-NOV-2000.
 PD 24-MAY-2000; 2000WO-US14296.
 XX 24-MAY-1999; 99US-0135631.
 PF (MAYO-) MAYO MEDICAL VENTURES.
 PR Federspiel MJ;
 PA WPI; 2001-032041/04.
 PI Inhibiting or preventing infectious agent transmission in mammalian
 DR transplant recipients, by introducing recombinant DNA comprising DNA
 XX encoding extracellular proteins of the agent into donor cells, such as
 PT swine cells -
 PT
 XX Claim 16; Page 109-111; 144pp; English.
 PS The present invention provides a method to prevent the transmission of
 XX infectious agents during xenotransplantation. This involves introducing
 CC to donor swine cells a recombinant DNA encoding a peptide fragment from
 CC the infectious agent, and then introducing these cells into the
 CC transplant recipient.
 XX
 SQ Sequence 4918 BP; 1382 A; 1173 C; 1183 G; 1180 T; 0 other;

alignment_scores:

Quality: 2549.50 Length: 669
 Ratio: 4.411 Gaps: 9
 Percent Similarity: 86.398 Percent Identity: 71.001

alignment_block:

US-09-171-553b-10 x AAC67022 ..
 Align seg 1/1 to: AAC67022 from: 1 to: 4918
 1 MetHisProThrLeuSerTrpArgHisLeuProThrArgGlyGlyGluPr 17
 |||||
 2187 ATGCATCCACGTTAAAGCGCGCCACCTCCCGATTCCGGGTGGAAAGCC 2236
 17 oLysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeu 34
 |||||
 2237 GAAAGACTGAAATCCCTTAAAGCTTGCGCTCCATCCGCTGGTTCCTTA 2286
 34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50
 |||||
 2287 CTCGTCAATAACTCTCAAGTTAATGGTAACGCCCTTGGGACAGCCCG 2336
 51 AsnProHisArgProLeuSerProThrTrpLeuIleIleAspProAspTh 67
 |||||
 2337 AACTCCCAATAACCCCTTATCTCTCACCTGGTTACTTACTGACTCCGGTAC 2386
 67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84
 |||||
 2387 AGGTATTAAATTAACAGCACTCAAGGGGAGGCTCCCTTGGGACCTGGT 2436
 84 rpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaVal. 98
 |||||
 2437 GGCCTGAATTATATCTGCTTCGATCAGTAATCCCTGGTCTCAATGAC 2486
 99 LysSerThrProProAsnLeuValArgSerTrpGlyPheTrpCysCysPr 115
 |||||
 2487 CAGGCCACACCCCGGATGACTCCGTCCTACGGGTTTACCTTTGCC 2536
 115 ogly. ...ThrCluLysGluLysTrpCysGlyGlySerGlyGluSerPheC 131
 |||||
 2537 AGGACCCCAATAATGAAGAATATTGTGAAATTCCTCAGGATTTCTTTT 2586

131 ySArgArgTrpSerCysValThrSerAsnAspGlyAspTrpLysTrpPro 147
 |||||
 2587 GCAAGCAATGAGCTGCTAACTTCTAATGATGGAAATGGAAATGGCCA 2636
 148 IleSerLeuGlnAspArgValLysPheSerPheValAsnSer. 161
 ::||| |||||
 2637 GTCTCTCAGCAAGACAGTAAGTTACTCTTTGTTTAAACAATCCTACCAG 2686
 162GlyProGlyLysTrpLys.MetMetL 170
 ||| |||||
 2687 TTATAATCAATTTAATTTATGCCATGGAGATGAAAGATTGGCAACAGC 2736
 170 ySLeuTrpLysAspLys.SerCysSerProSerAsp 181
 ::||| |||||
 2737 GGGTACAAAAGATGTACGAAATAAGCAATAAGCTGTCAATTCGTTAGAC 2786
 182 LeuAspTrpLeuLysIleSerPheThrGluLysGlyLysGlnGluAsnIl 198
 |||||
 2787 CTAGATTACTTAAATAAGTTTCACTGAAAAAGGAAAAACAAGAAATAT 2836
 198 eGlnLysTrpIleAsnGlyMetSerTrpGlyIleValPheTrpLysTrp 215
 |||||
 2837 TCAAAAGTGGTAATGGTATGCTTTGGGAATAGTGTAC.TATG 2880
 215 lyGlyGlyAla.GlySerThrLeuThrIleArgLeuArgIle 228
 |||||
 2881 GAGGCTCTGGGAGAAAGAAAGATGCTGTCTGACTATTCGCCCTCAGAATA 2930
 229 GluThrGlyThrGluProProValAlaMetGlyProAspLysValLeuAl 245
 |||||
 2931 GAAACTCAGATGGAACTCCGGTTGCTATAGGACCAATAAGGGTTGGC 2980
 245 aGluGlnGlyProProAlaLeuGluProProHisAsnLeuProValPro 262
 |||||
 2981 CGAACAGGACCTCCCAATCCAGAACAG. 3008
 262 lnLeuThrSerLeuArgProAspIleThrClnProProSerAsnSerThr 278
 |||||
 3009AGGCCATCTCTCAACCCCTCTGATTACATAACAACC 3044
 279 ThrGlyLeuIleProThrAsnThrProArgAsnSerProGlyValProVa 295
 ::|||
 3045 TCTGGTACATCCCACT.GAGCCTAACATCACTAT 3079
 295 LysThrGlyGlnArgLeuPheSerLeuIleGlnGlyAlaPheGlnAlaI 312
 |||||
 3080 TAAACAGGGGGGAAACTTTTAAACCTCATCCAGGAGGCTTTTCAAGCTC 3129
 312 leAsnSerThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSer 328
 |||||
 3130 TTAACCTCCAGACTCCAGAGGCTACCTCTCTTTGTTGGCTTTGCTTAGCT 3179
 329 SerGlyProProTrpTrpGluGlyMetAlaLysGluArgLysPheAsnVa 345
 |||||
 3180 TCGGGCCCACTTACTATCAGGGAATGGCTAGAGGAGGGAATTCATGT 3229
 345 lThrLysGluHisArgAsnGlnCysThrTrpGlySerArgAsnLysLeu 362
 |||||
 3230 GACAAAGGAACATAGAGACCAATGTACATGGGATCCCAAAATAAGCTTA 3279
 362 hrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysAlaProPro 378
 |||||
 3280 CCTTACTGAGGTTCTGGAAAGGCACCTGCATAGGATGGTTCCTCCCA 3329
 379 SerHisGlnHisLeuCysTrpSerThrValValTrpGluGlnAlaSerGl 395
 |||||
 3330 TCCCAACCAACCTTTGTAAACACACACTGAAGCCTTAACTCAACCTCGA 3379
 395 uAsnGlnTrpLeuValProGlyTrpAsnArgTrpTrpAlaCysAsnThrG 412
 |||||
 3380 GAGTCAGTATCTGGTACCTGGTTATCAGAGGTGGTGGCATGTAATACTG 3429
 412 lLeuThrProCysValSerThrSerValPheAsnGlnSerLysAspLeu 428

```

|||||
3430 GATTAACCCCTGTGTTCACCTTGGTTTCAACCAACTAAAGACTTT 3479
429 CysValMetValGlnIleValProArgValTyrHisProGluGluVa 445
|||||
3480 TGCCTTATGTCCTCAAAATGTCCCGGGGTGTACTACTATCCGAAAAGC 3529
445 lValLeuAspGluTyrAspTyrArgTyrAsnArgProLysArgGluProV 462
|||||
3530 ACTCCTTGATGAATATGACATATAGTATATCGGCCAAAGAGAGCCCA 3579
462 alSerLeuThrLeuAlaValMetLeuGlyLeuGlyThrAlaValGlyVal 478
|||||
3580 TATCCCTGACACTAGCTAGCTGTAATGCTCGGATTGGAGTGGCTGCAGGCGTG 3629
479 GlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeuGluLysG 495
|||||
3630 GGAACAGGAAGCGGTGCCCTAATCACAGAGACCCGCAACAGCTGGAGAAGG 3679
495 yLeuGlyGluLeuHisAlaAlaMetThrGluAspLeuArgAlaLeuLysG 512
|||||
3680 ACTTAGTAACCTACATCGAATTGTAACGGAAGATCTCCAAGCCCTAGAAA 3729
512 luSerValSerAsnLeuGluGluSerLeuThrSerLeuSerGluValVal 528
|||||
3730 AATCTGTCACTAGTAACCTGAGGAATCCCTAACCTCTTATCTGAAGTGGTT 3779
529 LeuGlnAsnArgArgGlyLeuAspLeuPheLeuArgGluGlyLe 545
|||||
3780 CTACAGAACAGAGGGGTTAGATCTCTTATTCTAAAGAGGAGGATT 3829
545 uCysAlaAlaLeuLysGluGluCysCysPheTyrValAspHisSerGlyA 562
|||||
3830 ATGTGTAGCCTTGAAGAGGAATGCTGTTTTATGTGGATCTTCAGGGG 3879
562 laIleArgAspSerMetAsnLysLeuArgLysLysLeuGluArgArg 578
|||||
3880 CCATCAGACTCCATCAACAAGCTTAGAGAAAGGTTGGAGAAAGCGTCA 3929
579 ArgGluArgGluAlaAspGlnGlyTyrPheGluGlyTyrPheAsnArgSe 595
|||||
3930 AGGAAAAGAAACACTACTCAAGGTTGGTTTCAGGGATGGTTCAACAGGTC 3979
595 rProTyrMetThrThrLeuLeuSerAlaLeuThrGlyProLeuValVal 612
|||||
3980 TCCTTGGTGGCTACCTTCTGCTTTAACAGGACCCCTTAATAGTCC 4029
612 euLeuLeuLeuLeuThrValGlyProCysLeuIleAsnArgPheValAla 628
|||||
4030 TCCTCTCTGTACTCACAGTTGGGCCATGTATTATTAAACAAGTTAATTGCC 4079
629 PheValArgGluArgValSerAlaValGlnIleMetValLeuArgGlnG 645
|||||
4080 TTCATTAGAACGAATTAAGTGCAGTCCAGATCATGTGTTACTTACAGAAC 4129
645 nTyrGln 647
|||||
4130 GTACCAA 4136

```

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT: AAC67018

seq_documentation_block:

ID AAC67018 standard; DNA; 1980 BP.

XX AAC67018;

XX AC AAC67018;

XX 27-MAR-2001 (first entry)

XX PERV-1-15 env protein coding sequence SEQ ID NO: 18.

XX Xenotransplantation; infectious agent; vaccine; ds.

XX Porcine endogenous retrovirus.

```

XX WO200071726-AL.
PN 30-NOV-2000.
XX 24-MAY-2000; 2000WO-US14296.
PF 24-MAY-1999; 99US-0135631.
XX (MAYO-) MAYO MEDICAL VENTURES.
PA Federspiel MJ;
XX WPI; 2001-032041/04.
XX
XX Inhibiting or preventing infectious agent transmission in mammalian
transplant recipients, by introducing recombinant DNA comprising DNA
PT encoding extracellular proteins of the agent into donor cells, such as
PT swine cells -
XX
XX Claim 16; Page 100-101; 144pp; English.
XX
XX The present invention provides a method to prevent the transmission of
infectious agents during xenotransplantation. This involves introducing
CC to donor swine cells a recombinant DNA encoding a peptide fragment from
CC the infectious agent, and then introducing these cells into the
CC transplant recipient.
XX
XX Sequence 1980 BP; 582 A; 452 C; 447 G; 499 T; 0 other;
SQ

```

alignment_scores:

Quality: 2546.50 Length: 667
 Ratio: 4.413 Gaps: 8
 Percent Similarity: 86.507 Percent Identity: 70.915

alignment_block:

US-09-171-553B-10 x AAC67018

Align seg 1/1 to: AAC67018 from: 1 to: 1980

```

1 MetHisProThrLeuSerTrpArgHisLeuProThrArgGlyGlyGluPr 17
|||||
1 ATGCATCCACGCTTAAGCGCGGCCACCTCCCGATTCGGGGTGAAGGCC 50
17 oLysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuT 34
|||||
51 GAAAGACTGAAATCCCTTAAGCTTCGCCCTCCATCGCGTTCCTTA 100
34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50
|||||
101 CTCTGTCAATAACTCTCAAGTTAATGGTAAACGCCCTTGTGACACCGCC 150
51 AsnProHisArgProLeuSerProThrTrpLeuIleAspProAspTh 67
|||||
151 AACTCCCATAAACCCCTTATCTCTACCTGGTTACTTACTGACTCCGGTAC 200
67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrpT 84
|||||
201 AGTATTATATATTAACAGCACTCAAGGGGAGGCTCCCTTGGGGACCTGTT 250
84 rpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaVal..... 98
251 GGCCTGAATTATATGTCCTCGCTTCGATCATCCCTGGTCTCAATGAC 300
99 LysSerThrProAsnLeuValArgSerTyrGlyPheTyrCysCysPr 115
|||||
301 CAGGCCACACCCCGGATGCTACTCGGTGCTTACGGGTTTACGTTTGCC 350
115 oGly...ThrGluLysGluLysTyrCysGlySerGlyGluSerPheC 131
|||||
351 AGGGCCCCCAATTAATGAAGAATATTGTGAAATCTCAGGATTCTTTT 400

```

```

131 ysArgArgTrpSerCysValThrSerAsnAspGlyAspTrpLysTrpPro 147
|||||
401 GCAAGCAATGGAGCTCGTAACCTCTAATGATGGGAATGGAAATGGCCA 450
|||||
148 IleSerLeuGlnAspArgValLysPheSerPheValAsnSer..... 161
|||||
451 GTCTCTCAGCAAGACAGAGTAAGTTACTCTTTTGTAAACAATCCTACCAG 500
|||||
162 .....GlyProGlyLysTrpLys.....MetMetL 170
|||||
501 TTATAATCAATTAATATATGGCCATGGGAGATGGAAAGATTGGCAACAGC 550
|||||
170 ysLeuTrpLysAspLys.....SerCysSerProSerAsp 181
|||||
551 GGGTACAAAAGATGTACCAAAATAGCAATTAAGCTGCTATTCGTTAGAC 600
|||||
182 LeuAspTrpLeuLysIleSerPheThrGluLysGlyLysGlnGluAsnI 198
|||||
601 CTAGATTACTTAAAAATAAGTTTCACTGAAAAGGAAAGCAAGAAATAT 650
|||||
198 eGlnLysTrpIleAsnGlyMetSerTrpGlyIleValPheTrpLysTrpG 215
|||||
651 TCAAAAGTGGGTAAATGGTATCTCTGGGGAATAGTGTACTATAGAGCT 700
|||||
215 lyGly...GlyAlaGlySerThrLeuThrIleArgLeuArgIleGluThr 230
|||||
701 CTGGGGAAGAAGAGATCTGTCTGATCTATTCGCCCTCAGAAATAGAACT 750
|||||
231 GlyThrGluProProValAlaMetGlyProAspLysValLeuAlaGluG 247
|||||
751 CAGATGGAAACCTCCGGTTGCTATAGACCAAAATAAGGGTTTGGCGGAACA 800
|||||
247 nGlyProProAlaLeuGluProProHisAsnLeuProValProGlnLeu 264
|||||
801 AGGACCTCCAAATCCAAAGAACAG..... 822
|||||
264 hrSerLeuArgProAspIleThrGlnProProSerAsnSerThrGly 280
|||||
823 .....AGGCATCTCTTAACCCCTCTGATTCAATACACCTCTCGGA 864
|||||
281 LeuIleProThrAsnThrProArgAsnSerProGlyValProValLysTh 297
|||||
865 TCAGTCCGCCACT.....GAGCCTAACATCACTATTAAAAAC 899
|||||
297 rGlyGlnArgLeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnS 314
|||||
900 AGGGCGGAACCTTTTAACTTCACTCCAGGGAGCTTTTCAAGCTCTTAACT 949
|||||
314 erThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSerSerGly 330
|||||
950 CCACGACTCCAGAGCTACCTCTCTCTGCTGGCTTGGCTAGCTTCGGGC 999
|||||
331 ProProTrpTrpGluGlyMetAlaLysGluArgLysPheAsnValThrLy 347
|||||
1000 CCACCTTACTATGAGGGAATGGCTAGAGGAGGAAATTCATATGTGACAAA 1049
|||||
347 sGluHisArgAsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeu 364
|||||
1050 GGAACATAGACCACTTATGATGGGATCCCAAAATAAGCTTACCCCTTA 1099
|||||
364 hrGluValSerGlyLysGlyThrCysIleGlyLysAlaProProSerHis 380
|||||
1100 CTGAGGTTTCTGGAAAAGGCACCTGCATAGGATGGTTCCTCCCATCCAC 1149
|||||
381 GlnHisLeuCysTrpSerThrValValTrpGluGlnAlaSerGluAsnG 397
|||||
1150 CAACACCTTTGTCACACACATCGAAGCCTTTAATCGAACCTCTGAGAGTCA 1199
|||||
397 nTrpLeuValProGlyTrpAsnArgTrpTrpAlaCysAsnThrGlyLeu 414
|||||
1200 GTAATCTGCTGCTGTTATGACAGTGTGGGATGTAATCTGATATAA 1249
|||||
414 hrProCysValSerThrSerValPheAsnGlnSerLysAspLeuCysVal 430
|||||

```

```

1250 CCCCTTGGTTTCCACCTTGGTTTTCACCAAACTAAAGACTTTTGGCGT 1299
|||||
431 MetValGlnIleValProArgValTrpHisProGluGluValValle 447
|||||
1300 ATGTCCTCAAAATTTGCCCGGGTGTACTACTATCCCGAAAAAGACGTCCT 1349
|||||
447 uAspGluTrpAspTrpArgTrpAsnArgProLysArgGluProValSerL 464
|||||
1350 TGATGAATATGACTATAGATATAATCGCCAAAAAGAGAGCCCATATCCC 1399
|||||
464 euThrLeuAlaValMetLeuGlyLeuGlyThrAlaValGlyValGlyThr 480
|||||
1400 TGACACATAGCTGTAATGCTCGATTTGGAGTGGCTGCAGCGCTGGGAACA 1449
|||||
481 GlyThrAlaAlaLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuG 497
|||||
1450 GGAAGGCTGCCCTTAATCACAGGACCCCAACAGCTGGAGAAAGACTTAG 1499
|||||
497 yGluLeuHisAlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerV 514
|||||
1500 TAACCTACATCGAATTGTAACGGAAAAATCTCCAAGCCCTAGAAAAATCTG 1549
|||||
514 alSerAsnLeuGluSerLeuThrSerLeuSerGluValValLeuGln 530
|||||
1550 TCAGTAACCTGGAGGAATCCCTAACCTCTTATCTGAAGTGGTCTACAG 1599
|||||
531 AsnArgArgGlyLeuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAl 547
|||||
1600 AACAGAAGGGGTAGATCTGTTATTTCTAAAGAGGAGGATATGTGT 1649
|||||
547 aAlaLeuLysGluGluCysCysPheTrpValAspHisSerGlyAlaIleA 564
|||||
1650 AGCCTTAAGAGGAGGATGCTGTTTTTATGTGGATCATTCAGGGGCCATCA 1699
|||||
564 rAspSerMetAsnLysLeuArgLysLysLeuGluArgArgArgGlu 580
|||||
1700 GAGACTCCATGAACAAGCTTAGAGAAGGTTGGAGAGCGCTCGAAGGAA 1749
|||||
581 ArgGluAlaAspGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTr 597
|||||
1750 AAGGAAACTACTCAAGGGTGGTTTGGAGGATGTTCAACAGGGTCTCCTTG 1799
|||||
597 pMetThrThrLeuLeuSerAlaLeuThrGlyProLeuValValLeuLeu 614
|||||
1800 GTTGGCTACCTCTTCTCTTAAACAGGACCTTAAATAGTCTCTCTCC 1849
|||||
614 euLeuLeuThrValGlyProCysLeuIleAsnArgPheValAlaPheVal 630
|||||
1850 TGTACTCAGCTGGGCCATGTATTATTAAACAAGTTAATTCCTTCATT 1899
|||||
631 ArgGluArgValSerAlaValGlnIleMetValLeuArgGlnGlnTrpG 647
|||||
1900 AGAGAAGAATAAGTGCAGTCCAGATCATGGTACTTAGACAACAGTACCA 1949
|||||
647 n 647
1950 A 1950

```

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAF77726

seq_documentation_block:

ID AAF77726 standard; DNA; 7333 BP.

XX AAF77726;

XX 23-MAY-2001 (first entry)

XX Defective retroviral genome isolated from PK-15 cell line.

XX Retrovirus; graft transplantation; xenotransplantation; PK-15 cell line;

XX ss.


```

5949 GTACCTGGTATTGACAGAGGGTGGCGATGTAATACTGGATTAAACCCCTTG 5998
416 sValSerThrSerValPheAsnGlnSerLysAspLeuCysValMetValG 433
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
5999 TGTTCACACCTTGGTTCACCAACAACTAAAGACTTTTGGCTTATGGTCC 6048
433 lnIleValProArgValTyrTyrHisProGluGluValValLeuAspGlu 449
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
6049 AAATTGTCCCGGGGTGTACTACTATCCCGAAAAAGCAGCTCTTGATGAA 6098
450 TyrAspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLe 466
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
6099 TATGACTATAGATATAATCGGCCAAAGAGAGAGCCCATATCCCTGACACT 6148
466 uAlaValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrA 483
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
6149 AGCTGTAATGCTCGGATTGGAGTGGCTCCAGGCGTGGGAACAGGACGG 6198
483 laAlaLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuGlyGluLeu 499
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
6199 CTGCCCTAATCAGAGCGCGCAACAGCTGGAGAAAGGACTTAGTAACCTA 6248
500 HisAlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAs 516
|||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
6249 CATCGAATTGTAAACGGAAGATCTCCAGCCCTAGAAAATCTGTGAGTAA 6298
516 nLeuGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgA 533
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
6299 CTGGAGGAATCCCTAACTCCTTATCTGAAGTGGTCTCAGAACACAGAA 6348
533 rgGlyLeuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaAlaLeu 549
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
6349 GGGGTTAGATCTGTATTCTTAAAGAAAGGAGGCGTTCGATGAGCCTTA 6398
550 LysGluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSe 566
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
6399 AAAGAGGAATGCTGCTTATGTAGATCACTCAGGAGCCATCAGAGACTC 6448
566 rMetAsnLysLeuArgLysLysLeuGluArgArgArgGluArgGluA 583
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
6449 CATGAGCAAGCTTAGAGAAAGGTTAGAGAGGCGTCGAAGGGAAGAGAGG 6498
583 laAspGlnGlyTrpPheGluGlyTyrPheAsnArgSerProTrpMetThr 599
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
6499 CTGACAGGGGTGTTGAAGATGGTTCACAGGTCTCCTGTGATGACC 6548
600 ThrLeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeu 616
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
6549 ACCCTGCTTCTGCTCAGCGGGCCCTAGTAGTCTGCTGCTGCTTACT 6598
616 uThrValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluA 633
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
6599 TACAGTTGGGCTCTGCTTAATTAATAGGTTGTTGCTTGTAGAGAAC 6648
633 rgValSerAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeu 649
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
6649 GAGTAGTGCACTCCAGATCATGTGACTTAGGCAACACTACCAAGGCCCTT 6698
650 LeuSerGlnGlyGluThrAspLeu 657
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
6699 CTGAGCAAGGAGAACTGACCTC 6722
seq_name: /SID52/gcgdata/geneseq/geneseq/NA197.DAT:AA74883
seq_documentation_block:
ID AA74883 standard; cDNA: 7393 BP.
XX
AC AA74883;
XX
DT 09-FEB-1998 (first entry)
XX
DE Porcine retrovirus cDNA (defective).
XX

```

```

KW Retrovirus; porcine; GAG protein; POL protein; ENV protein;
KW xenotransplantation; infectious; provirus; organ transplant; donor;
XX activated virus; PCR; ss.
OS Porcine retrovirus.
XX
FH Key Location/Qualifiers
FT CDS 598..2172
FT /tag= a
FT mat_peptide 598..2169
FT /tag= b
FT /note= "putative GAG protein"
FT CDS 2320..4737
FT /tag= c
FT /note= "putative POL coding region (partial) as
FT described in the specification"
FT mat_peptide 2320..3522
FT /tag= d
FT /note= "putative POL protein (partial)"
FT mat_peptide 3516..4328
FT /tag= e
FT /note= "putative POL protein (partial)"
FT mat_peptide 4332..4748
FT /tag= f
FT CDS 4738..6725
FT /tag= g
FT /note= "putative ENV coding region (partial) as
FT described in the specification"
FT mat_peptide 4752..6722
FT /tag= h
FT /note= "ENV protein (partial)"
XX
PN WO9721836-A1.
XX 19-JUN-1997.
XX
XX 13-DEC-1996; 96WO-US19680.
XX
XX 14-DEC-1995; 95OS-0572645.
XX (GEO ) GEN HOSPITAL CORP.
XX Fishman JA;
XX WPI; 1997-332804/30.
XX P-PSDB; AAW32091-W32095.
XX
XX New nucleic acid from porcine retroviruses - used for detecting
XX viruses in transplant or other tissue and for assessing risk of
XX transmitting infection to graft recipient
XX
XX Claim 16; Fig 2; 128pp; English.
XX
XX This cDNA sequence represents a defective purified swine retrovirus
XX found in PK-15 cells containing the putative coding regions for viral
XX GAG, POL and ENV proteins. There are a few in frame stop codons and
XX apparent frame shifts in the given coding sequence which alter features
XX of the translation. This sequence and PCR fragments generated from the
XX presence of porcine retroviruses prior to xenotransplantation.
XX Transplantation can increase the likelihood of retroviral activation if
XX intact and infectious proviruses are present. The porcine retroviral
XX sequence can be used to generate probes to determine the level (e.g.
XX copy number) of intact (i.e. potentially replicating) porcine provirus
XX sequences in a strain of xenograft transplantation donors. It can be
XX used to detect mutations, genetic lesions or viral recombinants and to
XX determine the histological localisation of activated retroviruses. Using
XX Polymerase Chain Reaction DNA Quantitation (PQO) on blood mononuclear
XX cells, infectivity titration and susceptibility testing can be
XX performed. Ultimately animal donors without intact porcine retroviral
XX sequences or a lower copy number of viral elements could be selected.
XX
XX Sequence 7393 BP; 2004 A; 1801 C; 1896 G; 1692 T; 0 other;
SQ

```

alignment_scores:

Quality: 2537.50 Length: 658
 Ratio: 4.452 Gaps: 9
 Percent Similarity: 86.626 Percent Identity: 72.796

alignment_block:

US-09-171-553B-10 x AAT74883 ..

Align seg 1/1 to: AAT74883 from: 1 to: 7393

22 lleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuThrIleTh 38
 4806 ATCCCTTAAGCTTCGCCCTCCATCGCTGGTTCCTTACTCTGTCAATAAC 4855
 38 rProGlnAlaSerSerLysArgLeuIleAspSerSerAsnProHisArgp 55
 4856 TCCTCAAGTTAATGGTAACGCCCTTGTGGACAGCCCGCAACTCCCATAAAC 4905
 55 toLeuSerProThrTrpLeuIleIleAspProAspThrGlyValThrVal 71
 4906 CCTTATCTCTACCTGGTTACTTACTGACTCCGGTACAGGATTAATATT 4955
 72 AsnSerThrArgGlyValAlaProArgGlyThrTrpTrpProGluLeuHi 88
 4956 AACAGCACTCAAGGGGAGCTCCCTTGGGACCTGGTGGCCTGAATTATA 5005
 88 spheCysLeuArgLeuIleAsnProAlaVal.....LysSerThrProp 103
 5006 TGTCTGCCTTCGATCAGTAATCCCTCGTCTCAATCAGCAGCCACAGCCC 5055
 103 roAsnLeuValArgSerThrGlyPheThrCysCysProGly...ThrGlu 118
 5056 CGATGACTCCGCTTACCGGTTTACGTTTGCACGAGCCACCCCAAT 5105
 119 LysGluLysThrCysGlySerGlySerGlyLeuSerPheCysArgArgTrpSe 135
 5106 AATGAAGAATATTGGGAATCCTCAGGATTTCTTTTGCAAGCAATGGAG 5155
 135 rCysValThrSerAsnAspGlyAspTrpLysTrpProLysSerLeuGlnA 152
 5156 CTGCATAACTTCTAATGATGGGAATTGGAATGGCCAGTCTCTCAGCAAG 5205
 152 spArgValLysPheSerPheValAsnSer..... 161
 5206 ACAGATGAAGTTACTCTTTTGTACCAATCCTACCAAGTTAATCAATTT 5255
 162GlyProGlyLysThrLys.....MetMetLysLeuThrLysAs 174
 5256 AATTATGGCCATGGAGATGGAAGATTGGCAACAGCGGGGTACAAAAGA 5305
 174 pLys.....SerCysSerProSerAspLeuAspThrLeuL 186
 5306 TGTACGAAATAAGCAAAATAAGCTCTCATTCCTTAGACCTAGATTACTTAA 5355
 186 ysIleSerPheThrGluLysGlyLysGlnGluAsnIleGlnLysTrpIle 202
 5356 AAATAAGTTTCTACGAAAGGAAACAAAGAAATATTTCAAAAGTGGGTA 5405
 203 AsnGlyMetSerTrpGlyIleValPheThrLysThrGlyGlyAla... 218
 5406 AATGGTATATCTTGGGAATAGTGAC.....TATGGAGGCTCTGGGAG 5449
 219GlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrG 233
 5450 AAGAAAGAGATCTCTTCTGACTATTCCCTCAGAAATAGAAACTCAGATGG 5499
 233 luProProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyPro 249
 5500 AACCTCGGTTGCTATAGACCAAAATAAGGTTTGGCCGCAACAGGACCT 5549
 250 ProAlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLe 266

5550 CCAATCCAAGACAG..... 5564
 266 uArgProAspIleThrGlnProProSerAsnSerThrThrGlyLeuIleP 283
 5565 .AGCCATCTCCTAAACCCCTCTGATTACAATAACAACCTCTGGATCAGTCC 5613
 283 roThrAsnThrProArgAsnSerProGlyValProValLysThrGlyGln 299
 5614 CCACT.....GAGCCTAACATCACTATTATAAACAGGGGGG 5648
 300 ArgLeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAs 316
 5649 AAACCTTTTAGCCTCATCCAGGAGCTTTTCAAGCTCTTAACCTCCAGCAG 5698
 316 pProAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProPro 333
 5699 TCCAGAGGCTACCTCTCTTGTGGTCTTGTAGCTTCGGGCCACCTT 5748
 333 YrTrpGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHis 349
 5749 ACTATGAGGAATGGCTAGAGGAGGAATTCATGTGACAAAGGAACAT 5798
 350 ArgAsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluVa 366
 5799 AGAGACCAATGTATACATGGGGATCCCAAAATAAGCTTACCTTACTGAGGT 5848
 366 lSerGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisL 383
 5849 TTCTGAAAAAGGCACCTGCTAGGATGGTTCCCCCATCCCAACACACC 5898
 383 euCysTrpSerThrValValTrpGluGlnAlaSerGluAsnGlnTrpLeu 399
 5899 TTTGTAAACCACTGAAGCCTTTAATCGAACCTCTGAGAGTCAATATCTG 5948
 400 ValProGlyTrpAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCy 416
 5949 GTACCTGGTTATGACAGGTGGTGGGATGTAATACTGGATTAAACCCCTTG 5998
 416 sValSerThrSerValPheAsnGlnSerLysAspLeuCysValMetValG 433
 5999 TGTTTCCACCTTGGTTTTCACCAACAAATAAGACTTTTGGCTTATGGTCC 6048
 433 lntIleValProArgValTrpTrpHisProGluGluValValLeuAspGlu 449
 6049 AAATTTGCCCGGGGTGTACTACTATCCGGAAGAACAGCTCTCTGTATGAA 6098
 450 TyrAspTrpArgTrpAsnArgProLysArgGluProValSerLeuThrLe 466
 6099 TATGACTATAGATATATCGGCCAAAAGAGAGCCCATATCCTTGACACT 6148
 466 uAlaValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThra 483
 6149 AGCTGTAATGCTCGGATTGGGAGTGGCTCAGCGGTGGGAACAGGAACGG 6198
 483 laAlaLeuIleThrGlyProGlnGlnLeuLysGlyLeuGlyGluLeu 499
 6199 CTGCCCTTAATCAGAGCCGCAACAGCTGGAGAAAGGACTTAGTAACCTTA 6248
 500 HisAlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAs 516
 6249 CATCGAATTGTACGGGAAGATCTCCAGGCCCTAGAAAAATCTGTAGTAA 6298
 516 nLeuGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgA 533
 6299 CCTGGAGGAATCCCTAACCTCCTTATCTGAAGTGGTCTTACAGAACAGAA 6348
 533 rgGlyLeuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaLeu 549
 6349 GGGGGTTAGATCTGTATTCTTAAAGAAAGAGGAGGTTTGTGTAGCCTTA 6398
 550 LysGluGluCysCysPheThrValAspHisSerGlyAlaIleArgAspSe 566

6399 AAGAGGAATGCTGCTTCTATCTAGATCACTCAGGAGCCATCAGAGACTC 6448
566 rMetAsnLysLeuArgLysLysLeuGluArgArgArgArgGluArgGluA 583
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6449 CATGACCAAGCTTAGAGAAAGCTTAGAGAGGCGTCGAAGGAAAGAGAGG 6498
583 laAspGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThr 599
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6499 CTGACAGAGGGTGGTTGAAGATGGTTCAACAGGTCTCTTGGATGACC 6548
600 ThrLeuSerAlaLeuThrGlyProLeuValValLeuValLeuLeuLe 616
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6549 ACCCTGCTTCTGCTCTGACGGGGCCCTAGTAGCTCTCTCTTACT 6598
616 uThrValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluA 633
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6599 TACAGTTGGCCTTGCTTAATTAATAGGTTTGTGCTTTGTAGAGAAC 6648
633 rgValSerAlaValGlnIleMetValLeuArgGlnGlnThrGlnGlyLeu 649
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6649 GAGTGAAGTGCAGTCCAGATCATGGTACTTAGGCAACAGTACCAGGCCTT 6698
650 LeuSerGlnGlyGluThrAspLeu 657
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6699 CTGAGCCAGGAGAACTGACCTC 6722

seq_name: /SID2/9c9data/geneseq/geneseq/NA2001.DAT: AAC67023

seq_documentation_block:

ID AAC67023 standard; DNA; 7873 BP.

AC AAC67023;

DT 27-MAR-2001 (first entry)

DE PERV env protein coding sequence SEQ ID NO: 23.

EW Xenotransplantation; infectious agent; vaccine; ds.

XX Porcine endogenous retrovirus.

XX WO200071726-A1.

XX 30-NOV-2000.

XX 24-MAY-2000; 2000WO-US14296.

XX 24-MAY-1999; 99US-0135631.

XX (MAYO-) MAYO MEDICAL VENTURES.

XX Federspiel MJ;

XX WPI; 2001-032041/04.

PT Inhibiting or preventing infectious agent transmission in mammalian.
PT transplant recipients, by introducing recombinant DNA comprising DNA
PT encoding extracellular proteins of the agent into donor cells, such as
PT swine cells -

PS Claim 16; Page 112-115; 144pp; English.

CC The present invention provides a method to prevent the transmission of
CC infectious agents during xenotransplantation. This involves introducing
CC to donor swine cells a recombinant DNA encoding a peptide fragment from
CC the infectious agent, and then introducing these cells into the
CC transplant recipient.

XX Sequence 7873 BP; 2200 A; 1913 C; 1962 G; 1798 T; 0 other;

alignment_scores:

Quality: 2514.50 Length: 703

Ratio: 4.358 Gaps: 10
Percent Similarity: 82.077 Percent Identity: 67.425
alignment_block:
US-09-171-553B-10 x AAC67023 ..
Align seg 1/1 to: AAC67023 from: 1 to: 7873
1 MethHisProThrLeuSerTrpArgHisLeuProThrArgGlyGlyGluPr 17
5012 ATGCATCCACGTTAAGCGCGCGCACCTCCCGATTCCGGGTGGAAGCC 5061
17 oLysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuT 34
5062 GAAAGAGACTGAAATCCCTTAAGCTTCGCCTCCATCGCTGGTTCCTTA 5111
34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50
5112 CTCGTCAATAACACCTCAAGTTAAGTAAACGCCCTGTGGACAGCCCG 5161
51 AsnProHisArgProLeuSerProThrTrpLeuIleIleAspProAspTh 67
5162 AACTCCCATAAACCCCTTATCTCTACCTGGTTACTTACTCACTCCGCTAC 5211
67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrpT 84
5212 AGGTATTAAATTAACAGACACTCAAGGGAGGCTCCCTTGGGACCTGGT 5261
84 rpProGluLeuHisPheCysLysArgLeuIleAsnProAlaVal..... 98
5262 GGCCTGAATTATATGCTCGCTTCGATCAGTAATCCTCGTCTCAATGAC 5311
99 LysSerThrProProAsnLeuValArgSerTyrGlyPheTyrCysCysPr 115
5312 CAGGCCACACCCCGGATGACTCCGCTGCTTACGGGTTTACGTTTGCC 5361
115 oGly...ThrGluLysGluLysTyrCysGlyGlySerGlyGluSerPheC 131
5362 AGGACCCCAATAATGAAGAATATTGTGAAATCCTCAGATTCTTTT 5411
131 ysArgArgTrpSerCysValThrSerAsnAspGlyAspTrpLysTrpPro 147
5412 GCAAGCAATGGAGCTGCTAACTTCTAATGATGGAAATGGAAATGCCA 5461
148 IleSerLeuGlnAspArgValLysPheSerPheValAsnSer..... 161
5462 GTCTCTCAGCAAGACAGAGTAAGTTACTCTTTTGTAAACAATCCTACCAG 5511
162GlyProGlyLysTyrLys.....MetMetL 170
5512 TTATAATCAATTTAATTATGCCATGGGAGATGGAAAGATTGGCAACAGC 5561
170 ysLeuTyrLysAspLys.....SerCysSerProSerAsp 181
5562 GGGTACAAAAGATGTAGCAAAATAAGCAATAAGCTGTCTCATTTAGAC 5611
182 LeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGluAsnII 198
5612 CTAGATTACTTAAATAAAGTTTCACTGAAAAAGGAAACAAAGAAATAT 5661
198 eGlnLysTrpIleAsnGlyMetSerTrpGlyIleValPheTyrLysTrpG 215
5662 TCAAAAGTGGTAAATGATGTCTTGGGAATAGTAGTAC.....TATG 5705
215 lyGlyGlyAla.....GlySerThrLeuThrIleArgLeuArgIle 228
5706 GAGCCTCTGGGAGAGAAAGAGGATCTGTCTGACTATTGCGCTCAGAATA 5755
229 GluThrGlyThrGluProProValAlaMetGlyProAspLysValLeuAl 245
5756 GAAACTCAGATGGAAACCTCCGGTGTATAGGACCAATAAAGGTTTGGC 5805
245 aGluGlnGlyProProAlaLeuGluProProHisAsnLeuProValProG 262

```

|||||
5806 CGAACAGGACCTCCAATCCAGAACAG..... 5833
262 InLeuThrSerLeuArgProAspIleThrGlnProProSerAsnSerThr 278
|||||
5834 .....AGGCCATCTCTTAACCCCTCTGATTACAATACACCC 5869
279 ThrGlyLeuIleProThrAsnThrProArgAsnSerProGlyValProVa 295
|||||
5870 TCTGGATCAGTCCCACT.....GAGCCTAACACTACTAT 5904
295 IlyThrGlyGlnArgLeuPheSerLeuIleGlnGlyAlaPheGlnAlaI 312
|||||
5905 TAAACAGGGCGAAACTTTTAAACCTCATCCAGGGAGCTTTTCAAGCTC 5954
312 LeAsnSerThrAspProAspAlaThrSerSerCysTyrLeuCysLeuSer 328
|||||
5955 TTAACCTCACCACTCCAGAGCTACCTCTCTTCTGCTTGGCTTGTAGCT 6004
329 SerGlyProProTyrTyrGluGlyMetAlaLysGluArgLysPheAsnVa 345
|||||
6005 TCGGCCCAACCTTACTATAGGGAATGGCTAGAGGAGGAAATTCATGT 6054
345 lThrLysGluHisArgAsnGlnCysThrTyrGlySerArgAsnLysLeuT 362
|||||
6055 GACAAAGGAACATAGAGACCAATATACATGGGATCCCAAAATAAGCTTA 6104
362 hrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysAlaProPro 378
|||||
6105 CCCTTACTAGGTTCTGGAAAGGACCTGCATAGGGATGGTTCGCCCA 6154
379 SerHisGlnHisLeuCysTyrSerThrValValTyrGluGlnAlaSerG 395
|||||
6155 TCCCAACCAACACCTTTTAAACACACCTGAAGCCTTTTAAATCGAACCTCTGA 6204
395 uAsnGlnTyrLeuValProGlyTyrAsnArgTyrTrpAlaCysAsnThrG 412
|||||
6205 GAGTCAGTATCTGTGCTACCTGTTATGACAGAGTGTGGCATGTAAATCTG 6254
412 lyLeuThrProCysValSerThrValPheAsnGlnSerLysAspLeu 428
|||||
6255 GATTAAACCCCTGTGTTTCCACCTTGGTTTTCACCAACCAATGAAGACTTT 6304
429 CysValMetValGlnIleValProArgValTyrTyrHisProGluGluVa 445
|||||
6305 TGCCTTATGGTCCAAATGTCCCGCGGTGTACTACTATCCCGAAAAAGC 6354
445 lValLeuAspGluTyrAspTyrArgTyrAsnArgProLysArgGluProv 462
|||||
6355 AGTCCCTTGATCAATATACATATAGATATAATCGGCCCAAAAGAGAGCCCA 6404
462 alSerLeuThrLeuAlaValMetLeuGlyLeuGlyThrAlaValGlyVal 478
|||||
6405 TATCCCTGCACATAGCTGTATGCTCGAGTGGAGTGGCTGCAGCGGTG 6454
479 GlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeuGlyLysG 495
|||||
6455 GGAACAGGAACGGGTGCGCTTAAATCACAGGACCGCAACAGCTGGAGAAGG 6504
495 yLeuGlyGluLeuHisAlaAlaMetThrGluAspLeuArgAlaLeuLysG 512
|||||
6505 ACTTAGTAACCTACATCGAATTTAAGCGGAAGATCTCCAAAGCCCTAGAA 6554
512 luSerValSerAsnLeuGluGluSerLeuThrSerLeuSerGluValVal 528
|||||
6555 AATCTGTCACTACCTGGAGAAATCCCTACCTCTTATCTGAAGTGGTT 6604
529 LeuGlnAsnArgArgGlyLeuAspLeuLeuPheLeuArgGluGlyLe 545
|||||
6605 CTACAGAACAGAGGGGGTTAGATCTGTTATTTCTAAAGAGAGGAGATT 6654
545 uCysAlaAlaLeuLysGluCysPheTyrValAspHisSerGlyA 562
|||||

```

```

6655 ATGTGTAGCCTTGAAGGAGGAATGCTCTTTTATGTGGATCATTCAGGGG 6704
562 laIleArgaspSerMetAsnLys..... 569
|||||
6705 CCATCAGAGACTCCATGACACAGCTTAGAGAAAGGACTGGAGAGCCCGG 6754
569 ..... 569
6755 CGGGTCTCGAACAAACCCAGACAGGTTCCTGTTTCAATTAAAGAACTGTC 6804
570 .....LeuArgLysLysLeuGluArgArgA 578
|||||
6805 GAAGTAACCGCTGAGCTAAAGCCAGCTTAGAGAAAGGTTGGAGAACGTC 6854
578 rArgGluArgGluAlaAspGlnGlyTrpPheGluGlyTrpPheAsnArg 594
|||||
6855 GAAGGAAAGAGAAACTACTCAAGAGTGGTTTGAGGGATGGTTCAACAGG 6904
595 SerProTrpMetThrThrLeuLeuSerAlaLeuThrGlyProLeuValVa 611
|||||
6905 TCTCCTTGGTGGCTACCTACTCTTCTGCTTTAACAGGACCTTAACTAGT 6954
611 lLeuLeuLeuLeuThrValGlyProCysLeuIleAsnArgPheValA 628
|||||
6955 CCTCCTCTGTTACTCACAGTTGGGCCATGTATTATTAACAAGTTAATG 7004
628 laPheValArgGluArgValSerAlaValGlnIleMetValLeuArgGln 644
|||||
7005 CCTTCAATTAGAACGAATAAGTGCAGTCCAGATCATGTACTTACTAGACAA 7054
645 GlnTyrGln 647
|||||
7055 CAGTACCAA 7063

seq_name: /SID52/gcgdata/geneseq/geneseq/NA1997.DAT:AA74811
seq_documentation_block:
ID AA74811 standard; cDNA; 8060 BP.
XX
AC AA74811;
XX
DT 11-FEB-1998 (first entry)
XX
DE Porcine retrovirus Tsukuba-1 cDNA.
XX
KW Retrovirus; porcine; GAG protein; POL protein; ENV protein;
KW xenotransplantation; infectious; provirus; organ transplant; donor;
KW activated virus; Tsukuba-1; PCR; ss.
XX
OS Porcine retrovirus.
XX
FH Key Location/Qualifiers
FT CDS 86..2002
FT /tag= a
FT /note= "putative ENV protein"
FT CDS 3112..4686
FT /tag= b
FT /note= "putative GAG protein"
FT CDS 4871..8060
FT /tag= c
FT /note= "putative POL protein (partial)"
XX
PN W09721836-AL.
XX
PD 19-JUN-1997.
XX
PF 13-DEC-1996; 96WO-US19680.
XX
PR 14-DEC-1995; 95US-0572645.
XX
PA (GEHO ) GEN HOSPITAL CORP.
XX
PI Fishman JA;

```

XX WPI; 1997-332804/30.

XX New nucleic acid from porcine retroviruses - used for detecting

PT viruses in transplant or other tissue and for assessing risk of

PT transmitting infection to graft recipient

XX Claim 1; Fig 1; 128pp; English.

XX This sequence represents the purified porcine retroviral cDNA

CC sequence of Tsukuba-1 and contains the putative coding regions for viral

CC proteins GAG, POL and ENV. This sequence and PCR fragments generated

CC from the sequence (see AAT74812-174882) could be used to screen organs

CC for porcine retroviruses prior to xenotransplantation. Transplantation

CC can increase the likelihood of retroviral activation if intact and

CC infectious proviruses are present. The porcine retroviral sequence can be

CC used to generate probes to determine the level (e.g. copy number) of

CC intact (i.e. potentially replicating) porcine provirus sequences in a

CC strain of xenograft transplantation donors. It can be used to detect

CC mutations, genetic lesions or viral recombinants and to determine the

CC histological localisation of activated retrovirus. Using Polymerase Chain

CC Reaction DNA Quantitation (PQ) on blood mononuclear cells, infectivity

CC titration and susceptibility testing can be performed. Ultimately animal

CC donors without intact porcine retroviral sequences or with a lower copy

CC number of viral elements could be selected.

XX SQ Sequence 8060 BP; 2233 A; 1959 C; 2012 G; 1856 T; 0 other;

alignment_scores:

Quality:	2402.00	Length:	655
Ratio:	4.312	Gaps:	9
Percent Similarity:	85.038	Percent Identity:	68.702

alignment_block:

us-09-171-553b-10 x AAT74811 ..

Align seg 1/1 to: AAT74811 from: 1 to: 8060

1 MetHisProThrLeuSerTrpArgHisLeuProThrArgGlyGluPr 17

86 ATGATCCACCGTTAAGCGCGCCACCTCCCGATTCCGGGTGGAAAGCC 135

17 OlyArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeu 34

136 GAAAGACACTCAAAATCCCTTAAGCTTCGCCCTCCCATCGGTTCCTTA 185

34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuLeuAspSer 50

186 CTCTGCAATAACCTCTCAGACTAATGGTATCGGCATAGGACAGACGCTG 235

51 AsnProHisArgProLeuSerProThrTrpLeuIleAlaAspProAspTh 67

236 AACTCCCATAAACCTTATCTCTCACCCTGGTTAATTACTGACTCCGCGAC 285

67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84

286 AGGTATATATATACACACACACTCAAGGGAGGCTCTTTAGGAACCTGT 335

84 rProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100

336 GGCCTGATCTATACGTTTCCTTCAGATCAGTTATTCCTAGCTGACCTCA 385

101 ThrProAsnLeuValArgSerThrGlyPheThrCysCysProGlyTh 117

386 ...CCCCAGATATCCTCCATGCTCAGGATTTATGTTTGGCCAGGACC 432

117 rGluLysGlu...LysTyrCysGlyCysSerGlyGluSerPheCysArg 133

433 ACCAAATATGGAACATTCGGGAATCCCGAGAGATTCCTTTTGAAC 482

133 rGTPSerCysValThrSerAsnAspGlyAspTrpLysTrpProIleSer 149

483 AATGGAACCTGTGTAACCTCTAATGATGATATTTGGAATGGCCAACTCT 532

150 LeuGlnAspArgValLysPheSerPheValAsn.....SerGlyProGl 164

533 CAGCAGATAGGGTAAGTTTTTCTTATCTCAACACCTATATACAGCTG 582

164 yLysTyrLysMetMetLysLeuTyrLysAspLysSer.....CysSerP 179

583 ACAATTTAATTAATCCTGACCTGGATTAGAACTGGAAGCCCAAGTCTCT 632

179 roSerAspLeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGln 195

633 CTTGAGACCTAGATACCTAAATAAAGTTTCTAGAGAAAGGAACAA 682

196 GluAsnIleGlnIlystrPileAsnGlyMetSerTrpGlyIleValPheTy 212

683 GAAATATCTCTAAATGGTAAATGGTATGCTTGGGAATGGTATAT... 730

212 rLysTyrGlyGlyGlyAla.....GlySerThrLeuThrIleArgL 226

731 ...TATGGAGGCTCGGGTAAACAACACAGGCTCCATTTCTAACTATTGCC 776

226 euArgIleGluThrGlyThrGluProProValAlaMetGlyProAspLys 242

777 TCAAAATAAACAG...CTGAGGCTCCATGGCTATAGGACCAATACG 823

243 ValLeuAlaGluGlnGlyProProAlaLeuGluProProHisAsnLeuPr 259

824 GTCTTGACGGGTCAAGAGACCCCAACCAAGGACCA..... 859

259 oValProGlnLeuThrSerLeuArgProAspIleThrGlnProProSerA 276

860GGACCATCTCTCTA 872

276 snSerThrThrGlyLeuIleProThrAsnThrProArgAsnSerProGly 292

873 ACATAACTCTGGATCAGACCCCACTGAGCTAGCAGCAGACT..... 916

293 ValProValLysThrGlyGlnArgLeuPheSerLeuIleGlnGlyAlaPh 309

917AAATGGGGGCAAACTTTTAGCCTCATCCAGGAGCTTT 957

309 eGlnAlaIleAsnSerThrAspProAspAlaThrSerSerCysTrpLeuC 326

958 TCAAGCTCTTAACCTCCAGACTCCAGAGGCTACCTCTCTGTGTGGCTAT 1007

326 ysLeuSerSerGlyProProTyrTyrGluGlyMetAlaLysGluArgLys 342

1008 GCTTAGCTTTGGGCCCACCTTACTATCAAGGAATGGCTAGAGAGGAAA 1057

343 PheAsnValThrLysGluHisArgAsnGlnCysThrTrpGlySerArgAs 359

1058 TTCAATGTGACAAAAGAACATAGAGACCAATGCACATGGGATCCCAAAA 1107

359 nLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysA 376

1108 TAAAGCTTACCTTACTAGGTTCTGGAAGAAGGCCTGCTAGGAAAGG 1157

376 laProProSerHisGlnHisLeuCysTyrSerThrValValTyrGluGln 392

1158 TTCCCATCCCAACACCTTTGTAACCACTGAAGCCTTTAATCAA 1207

393 AlaSerGluAsnGlnTyrLeuValProGlyTyrAsnArgTrpTrpAlaCy 409

1208 ACCCTCTGAAAGTCAATATCTGGTACCTGGTATGACAGGTTGGGGCATG 1257

409 sAsnThrGlyLeuThrProCysValSerThrSerValPheAsnGlnSerL 426

1258 TAATCTGGAATTAACCCCTGTGTTCCTCCACTTGGTTTAAACCAACTA 1307

426 ysAspLeuCysValMetValGlnIleValProArgValTyrTrpHisPro 442

1308 AAGATTTTTCATTTGTCCTCAAAATTTGTCCTCCAGTGTATTACTATCC 1357

326 ysLeuSerSerGlyProProTyrTyrGluGlyMetAlaLysGluArgLys 342
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6542 GCTTACGCTTCGGGCCACCTTACTATGAAGGAATGGCTAGAAAGAGGAAA 6591

343 PheAsnValThrLysGluHisArgAsnGlnCysThrTyrGlySerArgAs 359
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6592 TTCATGTGCACAAAGAACATAGACCANTGCACATGGGGATCCCAAAA 6641

359 nLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysA 376
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6642 TAAGCTTACCTTACTAGGTTTCTGGAAAAGGCACCTGCATAGGAAAGG 6691

376 laProProSerHisGlnHisLeuCysTyrSerThrValValTyrGluGln 392
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6692 TTCGCCCATCCCAACACACCTTTTGAACACCACTCAAGCCCTTAAATCAA 6741

393 AlaSerGluAsnGlnTyrLeuValProGlyTyrAsnArgTyrTyrAlaCys 409
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6742 ACCTCTGAGAGTCATATCTGTGTACTGTTATGACAGGTGGTGGCGATG 6791

409 sAsnThrGlyLeuThrProCysValSerThrSerValPheAsnGlnSerL 426
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6792 TAATACTGGATTAAACCCCTGTGTCTCCACCTTGGTTTAAACCAACTA 6841

426 ysAspLeuCysValMetValGlnIleValProArgValTyrThrHisPro 442
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6842 AAGATTTTTCGCAATATGGTCCAAATGTGTCCCGAGTGTATATCTATCCC 6891

443 GluGluValValLeuAspGluTyrAspTyrArgTyrAsnArgProLysAr 459
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6892 GAAAGACCAATCCTTGATGAATATGACTACAGAAATCATCGACAAAGAG 6941

459 gGluProValSerLeuThrLeuAlaValMetLeuGlyLeuGlyThrAlaV 476
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6942 AGAACCCATATCTCTACACCTTGTGTGTATGCTCGGACTTGGAGTGCAG 6991

476 aLcGlyValGlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeu 492
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6992 CAGGTGTAGAACAGAACAGCTGCCCTGGTCCAGGACCCACAGCAGCTA 7041

493 GluLysGlyLeuGlyGluLeuHisAlaAlaMetThrGluAspLeuArgAl 509
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
7042 GAAACAGGACTTAGTAACCTACATCGAATTTGTAACAGAAATCTCCAAGC 7091

509 aLeuLysGluSerValSerAsnLeuGluSerLeuThrSerLeuSerG 526
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
7092 CCTAGAAAATCTGTACGTAACCTGGAGGAATCCCTAACCTCTTATCTG 7141

526 luValValLeuGlnAsnArgArgGlyLeuAspLeuPheLeuArgGlu 542
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
7142 AGTAGTCTCAGAAATAGAGAGGTTAGATTATTTATTTCTAAAGAA 7191

543 GlyGlyLeuCysAlaAlaLeuLysGluGluCysCysPheTyrValAspH 559
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
7192 GGAGGATTATGTGTAGCTTGAAGGAGGAATGCTGTTTTATGTGGATCA 7241

559 sSerGlyAlaIleArgAspSerMetAsnLysLeuArgLysLysLeuGluA 576
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
7242 TTCAGGGGCCATCAGAGACTTCCATGAACAGCTTAGAAGAAAGGTGGAGA 7291

576 rArgArgArgGluArgGluAlaAspGlnGlyTyrPheGluGlyTyrPhe 592
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
7292 AGCGTCGAAGGAAAGAAAGAACTACTCAAGGGTGGTGGGGATGGTTC 7341

593 AsnArgSerProTyrMetThrThrLeuLeuSerAlaLeuThrGlyProLe 609
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
7342 AACAGGCTCTCTTGGTGGCTACCTTCTGCTTTTAAACAGGACCCCT 7391

609 uValValLeuLeuLeuThrValGlyProCysLeuLeuLeuAsnArgP 626
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
7392 AATAGTCCCTCCCTGTTACTCACAGTGGGCCATGTATTATTAACAAGT 7441

626 heValAlaPheValArgGluArgValSerAlaValGlnIleMetValLeu 642

7442 TAATTGCTTCAATTAGAACGAATAAGTGCAGTCCAGATCATGGTACTT 7491

643 ArgGlnGlnTyrGln 647

7492 AGACAACAGTACCAA 7506

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAF77727

seq_documentation_block:

ID AAF77727 standard; DNA; 8132 BP.

XX AAF77727;

XX 23-MAY-2001 (first entry)

XX Nucleotide sequence of a retrovirus found in miniature swine.

XX Retrovirus; graft transplantation; xenotransplantation; miniature swine;

XX Unidentified.

XX US6190861-B1.

XX 20-FEB-2001.

XX 13-DEC-1996; 96US-0766528.

XX 14-DEC-1995; 95US-0572645.

XX (GCHO) GEN HOSPITAL CORP.

XX Fishman JA;

XX WPI: 2001-256211/26.

XX P-PSDB; AAB73285, AAB73286, AAB73287.

XX Assessing risk of endogenous retroviruses in clinical practice and in xenotransplantation, comprises using probe sequences derived from swine or miniature swine retroviral genome -

XX Claim 1; Fig 3; 127pp; English.

XX The present invention relates to a method for screening a cell or tissue for the presence or expression of a retrovirus (RV), comprising contacting a target nucleic acid from the cell or tissue with a second nucleic acid from the present invention (e.g. the present sequence or a fragment thereof). The method is useful for RV detection and to assess graft transplantation risk. Screening of animals allows the elimination of donors with active replication of known viruses. Inactive proviruses can be detected and inactivated, allowing identification and elimination of potential human pathogens derived from swine in a manner not possible in the outbred human organ donor population and is important to the development of human xenotransplantation.

XX Sequence 8132 BP; 2248 A; 1977 C; 2037 G; 1870 T; 0 other;

alignment_scores:

Quality: 2395.00 Length: 655
Ratio: 4.292 Gaps: 10
Percent Similarity: 85.191 Percent Identity: 68.855

alignment_block:

US-09-171-553B-10 x AAF77727 ..

Align seg 1/1 to: AAF77727 from: 1 to: 8132

1 MethisProThrLeuSerTyrArgHisLeuProThrArgGlyGluPr 17

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
5620 ATGATCCACGTTAAACCGCGCCACCTCCCGATTCGGGGTGGAAAGCC 5669

```

17 oLysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuT 34
5670 GAAAAGACTGAAAAATCCCTTAAGCTTCGGCTCCATCGCGTGGTTCCTTA 5719
34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5720 CTCGTCAATAACCTCTCAGACTAATGGTATGGCATAGGACAGCGCTG 5769
51 AsnProHisArgProLeuSerProThrTrpLeuIleIleAspProAspTh 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5770 AACTCCCATAAACCTTATCTCACCTGGTTAATTACTGACTCCGGCAC 5819
67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrpT 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5820 AGGTATTAAATCAACAACACTCAAGGGGAGGCTCCTTTAGGAACCTGTT 5869
84 rpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5870 GGCCTGATCATAGCTTGGCTCAGATCAGTTATTCCTAGCTGACCTCA 5919
101 ThrProAsnLeuValArgSerThrGlyPheThrCysCysProGlyTh 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5920 ...CCCCAGATATCCCTCCATGCTCAGCGATTTTATGTTTGCCAGGACC 5966
117 rGluLysGlu...LysThrCysGlyGlySerGlyGluSerPheCysArgA 133
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5967 ACCAAATTAATGGAACAACTTCGGAATCCAGAGATTTCTTTGTAAC 6016
133 rgrPsrCysValThrSerAsnAspGlyAspTrpLysTrpProIleSer 149
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6017 AATGGAACTGTGAACCTCTAATGATGGATATTGGAATGGCCAACTCT 6066
150 LeuGlnAspArgValLysPheSerPheValAsn.....SerGlyProG 164
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6067 CAGCAGATAGGTAAAGTTTCTTATGTCACACCTATACAGCTCTGG 6116
164 yLysThrLysMetMetLysLeuThrLysAspLysSer.....CysSerP 179
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6117 ACAATTTAATACCTGACCTGGATTAGAACTGGAAGCCCAAGTGCTCTC 6166
179 roSerAspLeuAspThrLeuLysIleSerPheThrGluLysGlyLysGln 195
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6167 CTTCAGACCTTAGATTACCTAAAAATAAGTTTCTACTGAGAAAGGAAACAA 6216
196 GluAsnIleGlnLysTrpIleAsnGlyMetSerTrpGlyIleValPheTy 212
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6217 GAAATATCTCTAAATGGGTAAATGATGCTCTGGGGAATGGTATAT.. 6264
212 rLysThrGlyGlyGlyAla.....GlySerThrLeuThrIleArgL 226
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6265 ...TATGAGGCTCGGGTAAACAACACGAGCTCCATTCCTAATTCGCGC 6310
226 euArgIleGluThrGlyThrGluProProValAlaMetGlyProAspLys 242
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6311 TCAAAATAAACACAG...CTGGAGCTCCCAATGGCTATAGGACCAATACG 6357
243 ValLeuAlaGluGlnGlyProProAlaLeuGluProProHisAsnLeuPr 259
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6358 GTCTTGACGGGTCAAGACCCCAACCAAGGACCA..... 6393
259 oValProGlnLeuThrSerLeuArgProAspIleThrGlnProProSerA 276
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6394 ..... 6406
276 snSerThrThrGlyLeuIleProThrAsnThrProArgAsnSerProGly 292
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6407 ACATAACTCTTGATCAGACCCCACTGAGTCT.....AACAGC..... 6444
293 ValProValLysThrGlyGlnArgLeuPheSerLeuIleGlnClyAlaph 309
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6445 ...ACGACTAAATGGGGCAAACTTTTATGCTCATCCAGGAGCTTT 6491
309 eGlnAlaIleAsnSerThrAspProAspAlaThrSerSerCysTrpLeuC 326
|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6492 TCAAGCTCTTAACCTCACGACTCCAGAGGCTACCTCTTCTTGGCTAT 6541
326 yLysLeuSerSerGlyProProThrTrpGluGlyMetAlaLysGluArgLys 342
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6542 GCTTACGCTTCGGGGCCACCTTACTATGAAGGAATGGCTAGAAGGGGAAA 6591
343 PheAsnValThrLysGluHisArgAsnGlnCysThrTrpGlySerArgAs 359
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6592 TTCAATGTGACAAAAGACATAGAGACCAATGCACATGGGATCCCAAAA 6641
359 nLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysA 376
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6642 TAAGCTTACCTTACTGAGCTTCTGGAAGAGGACCTGCATAGAAAGG 6691
376 laProProSerHisGlnHisLeuCysTyrSerThrValValTyrGluGln 392
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6692 TTCCCCCATCCCAACCAACCTTTGTAAACCACTCAAGCCTTTAATCAA 6741
393 AlaSerGluAsnGlnTyrLeuValProGlyTyrAsnArgTrpTrpAlaCy 409
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6742 ACCTCTGAGAGTCAATATCTGTACTCTGTTATGACAGGTGGTGGCATG 6791
409 sAsnThrGlyLeuThrProCysValSerThrSerValPheAsnGlnSerL 426
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6792 TAATACTGGATTAAACCCCTTGTGTTCCACCTTGGTTTAAACCAACTA 6841
426 yAspLeuCysValMetValGlnIleValProArgValTyrTrpHisPro 442
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6842 AAGATTTTGCATATATGTTCCAAATGTTCCCGAGTGTATTACTATCCC 6891
443 GluGluValValLeuAspGlnTyrAspTyrArgTyrAsnArgProLysAr 459
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6892 GAAAAGCAATCCTTGATGAATATGACTACAGAAATCATCGACAAAAGAG 6941
459 gGluProValSerLeuThrLeuAlaValMetLeuGlyLeuGlyThrAlaV 476
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6942 ACAACCCATATCTCTGACACTTGCTGTGATGCTCGGACTTGGAGTGGCAG 6991
476 aLGLyValGlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeu 492
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6992 CAGGTGTAGGAACAGGAACAGCTGCCCTGGTCACGGGACACAGCAGTA 7041
493 GluLysGlyLeuGlyGluLeuHisAlaAlaMetThrGluAspLeuArgAl 509
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
7042 GAAACAGGACTTGTATACCTACATCGAATTTGTAACGAAGATCTCCAAGC 7091
509 aLeuLysGluSerValSerAsnLeuGluGluSerLeuThrSerLeuSerG 526
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
7092 CCTAGAAAATCTGTCTACCTAACCTGGAGGAATCCCTAACCTTATCTG 7141
526 luValValLeuGlnAsnArgArgGlyLeuAspLeuLeuPheLeuArgGlu 542
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
7142 AAGTAGTCTACAGAAATAGAAGGGTTAGATTATTATTATTCTAAAGAA 7191
543 GlyGlyLeuCysAlaAlaLeuLysGluGluCysCysPheTyrValAspHi 559
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
7192 GGAGATTATGTGTAGCCCTTGAAGGAGGAATGCTGTTTTATGTGGATCA 7241
559 sSerGlyAlaIleArgAspSerMetAsnLysLeuArgLysLysLeuGluA 576
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
7242 TTCAGGGCCCATCAGAGACTCCATGAACAAGCTTAGACAAAGTTGGAGA 7291
576 rgArgArgArgGluArgGluAlaAspGlnGlyTrpPheGluGlyTrpPhe 592
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
7292 AGCGTCGAAGGGGAAAAGGAAACTACTCAAGGGTGGTTTGGAGGATGGTTC 7341
593 AsnArgSerProTrpMetThrThrLeuLeuSerAlaLeuThrGlyProLe 609
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
7342 AACAGGTCTCTTGTGGTGGCTACCTTCTTCTTTAAGGAGCCCTT 7391
609 uValValLeuLeuLeuThrValGlyProCysLeuIleAsnArgP 626
|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2002, 03:05:27 ; Search time 624.99 Seconds
(without alignments)
27.435 Million cell updates/sec

Title: US-09-171-553B-13

Perfect score: 20

Sequence: 1 gatgctctctgccccttgg 20

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 930621 seqs, 428662619 residues 1861242

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_1101.*

1: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	AAV09704	Porcine retrovirus
2	20	100.0	18	AAV09698	Porcine retrovirus
3	20	100.0	18	AAV09698	Porcine retrovirus
4	20	100.0	18	AAV09698	Porcine retrovirus
5	20	100.0	18	AAV09698	Porcine retrovirus
6	20	100.0	18	AAV09698	Porcine retrovirus
7	18.4	92.0	3482	AAV82745	Porcine retrovirus
8	18.4	92.0	4402	AAV82745	Porcine retrovirus
9	18.4	92.0	4918	AAV82745	Porcine retrovirus
10	18.4	92.0	6076	AAV82745	Porcine retrovirus
11	18.4	92.0	6076	AAV82745	Porcine retrovirus

ALIGNMENTS

12	18.4	92.0	7362	22	AAV09704	Porcine retrovirus
13	18.4	92.0	7873	22	AAV09704	Porcine retrovirus
14	18.4	92.0	7892	18	AAV09704	Porcine retrovirus
15	18.4	92.0	8060	18	AAV09704	Porcine retrovirus
16	18.4	92.0	8060	22	AAV09704	Porcine retrovirus
17	18.4	92.0	8132	22	AAV09704	Porcine retrovirus
18	17.4	87.0	300	21	AAV09704	Porcine retrovirus
19	17.4	87.0	773	21	AAV09704	Porcine retrovirus
20	16.8	84.0	335	22	AAV09704	Porcine retrovirus
21	16.8	84.0	335	22	AAV09704	Porcine retrovirus
22	16.8	84.0	335	22	AAV09704	Porcine retrovirus
23	16.8	84.0	578	22	AAV09704	Porcine retrovirus
24	16.8	84.0	578	22	AAV09704	Porcine retrovirus
25	16.8	84.0	926	21	AAV09704	Porcine retrovirus
26	16.8	84.0	38186	22	AAV09704	Porcine retrovirus
27	16.8	84.0	38186	22	AAV09704	Porcine retrovirus
28	16.8	84.0	72928	21	AAV09704	Porcine retrovirus
29	16.8	84.0	72928	21	AAV09704	Porcine retrovirus
30	16.4	82.0	2331	21	AAV09704	Porcine retrovirus
31	16	80.0	23	18	AAV09704	Porcine retrovirus
32	16	80.0	23	18	AAV09704	Porcine retrovirus
33	16	80.0	23	22	AAV09704	Porcine retrovirus
34	16	80.0	23	22	AAV09704	Porcine retrovirus
35	16	80.0	2035	15	AAV09704	Porcine retrovirus
36	16	80.0	2401	19	AAV09704	Porcine retrovirus
37	15.8	79.0	519	21	AAV09704	Porcine retrovirus
38	15.8	79.0	1660	21	AAV09704	Porcine retrovirus
39	15.8	79.0	1667	21	AAV09704	Porcine retrovirus
40	15.8	79.0	10723	18	AAV09704	Porcine retrovirus
41	15.8	79.0	10723	18	AAV09704	Porcine retrovirus
42	15.8	79.0	13633	14	AAV09704	Porcine retrovirus
43	15.4	77.0	249	22	AAV09704	Porcine retrovirus
44	15.4	77.0	480	21	AAV09704	Porcine retrovirus
45	15.4	77.0	1153	13	AAV09704	Porcine retrovirus

AAV09704	standard; DNA: 20 BP.
AAV09704	
19-MAY-1998	(first entry)
Porcine retrovirus DNA PCR primer 3.	
Porcine retrovirus; POEV; POL protein; ENV protein; GAG protein; vaccine; diagnosis; xenotransplantation; prophylactic; therapeutic; PCR primer; ss.	
Synthetic.	
Porcine retrovirus.	
WO9740167-A1.	
30-OCT-1997.	
18-APR-1997.	97WO-GH01087.
10-FEB-1997.	97GB-0002668.
19-APR-1996.	96GB-0008164.
(IMUT-) IMUTRAN LTD.	
(ONE-) ONE BIOTECH LTD.	
Galbraith DN, Haworth C, Lees GM, Smith KT;	
WPI: 1997-535851/49.	
Polynucleotide encoding porcine retrovirus expression product -	

CC This DNA sequence encodes the porcine retrovirus (PoRV) polymerase (POL) and envelope (ENV) proteins. These proteins can be used to develop viral

[illegible]

QY 1 gatgagctctcctgaccttg 20
|||||

DB 4442 gatggtctctgccttgg 4461

RESULT 4
AA74883
ID AA74883 standard; cDNA: 7393 BP.
XX
AC AA74883;
XX
DT 09-FEB-1998 (first entry)
XX
DE Porcine retrovirus cDNA (defective).
XX
KW Retrovirus; porcine; GAG protein; POL protein; ENV protein;
KW xenotransplantation; infectious; provirus; organ transplant; donor;
XX activated virus; PCR; ss.
XX
OS Porcine retrovirus.
XX
FH Key Location/Qualifiers
FT CDS 598..2172
FT /*tag= a
FT 598..2169
FT mat_peptide
FT /*tag= b
FT /note= "putative GAG protein"
FT 2320..4737
FT /*tag= c
FT /note= "putative POL coding region (partial) as
FT described in the specification"
FT 2320..3522
FT mat_peptide
FT /*tag= d
FT /note= "putative POL protein (partial)"
FT 3516..4328
FT /*tag= e
FT /note= "putative POL protein (partial)"
FT 4332..4748
FT /*tag= f
FT 4738..6725
FT /*tag= g
FT /note= "putative ENV coding region (partial) as
FT described in the specification"
FT 4752..6722
FT /*tag= h
FT /note= "ENV protein (partial)"
FT
PN WO9721836-A1.
XX
PD 19-JUN-1997.
XX
PF 13-DEC-1996; 96WO-US19680.
XX
PR 14-DEC-1995; 95US-0572645.
XX
PA (GEHO) GEN HOSPITAL CORP.
XX
PI Fishman JA;
XX
DR WPI; 1997-332804/30.
DR P-PSDB; AAW32091-W32095.
XX
PT New nucleic acid from porcine retrovirus - used for detecting
PT viruses in transplant or other tissue and for assessing risk of
PT transmitting infection to graft recipient
XX
PS Claim 16; Fig 2; 128bp; English.
XX
CC This cDNA sequence represents a defective purified swine retrovirus
CC found in PK-15 cells containing the putative coding regions for viral
CC GAG, POL and ENV proteins. There are a few in frame stop codons and
CC apparent frame shifts in the given coding sequence which alter features
CC of the translation. This sequence and PCR fragments generated from the
CC sequence (see AA74812-774882) could be used to screen organs for the
CC presence of porcine retroviruses prior to xenotransplantation.

CC Transplantation can increase the likelihood of retroviral activation if
CC intact and infectious proviruses are present. The porcine retroviral
CC sequence can be used to generate probes to determine the level (e.g.,
CC copy number) of intact (i.e., potentially replicating) porcine provirus
CC sequences in a strain of xenograft transplantation donors. It can be
CC used to detect mutations, genetic lesions or viral recombinants and to
CC determine the histological localization of activated retroviruses. Using
CC Polymerase Chain Reaction DNA Quantitation (PDQ) on blood mononuclear
CC cells, infectivity titration and susceptibility testing can be
CC performed. Ultimately animal donors without intact porcine retroviral
CC sequences or a lower copy number of viral elements could be selected.
XX
SQ Sequence 7393 BP; 2004 A; 1801 C; 1896 G; 1692 T; 0 other.

Query Match 100.0%; Score 20; DB 18; Length 7393;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gatggtctctgccttgg 20
DB 4442 gatggtctctgccttgg 4461

RESULT 5
AAV09699
ID AAV09699 standard; DNA; 8196 BP.
XX
AC AAV09699;
XX
DT 19-MAY-1998 (first entry)
XX
DE Porcine retrovirus DNA encoding, GAG, POL and ENV.
XX
KW Porcine retrovirus; PoEV; POL protein; ENV protein; GAG protein;
KW vaccine; diagnosis; xenotransplantation; prophylactic; therapeutic; ds.
XX
OS Porcine retrovirus.
XX
FH Key Location/Qualifiers
FT CDS 576..2126
FT /*tag= a
FT /product= GAG protein
FT /note= "viral core polypeptide"
FT 2143..5733
FT /*tag= b
FT /product= POL protein
FT /note= "polymerase peptide sequence as given in
FT specification"
FT 5606..7576
FT /*tag= c
FT /product= ENV protein
FT /note= "envelope protein"
FT
PN WO9740167-A1.
XX
PD 30-OCT-1997.
XX
PF 18-APR-1997; 97WO-GB01087.
XX
PR 10-FEB-1997; 97GB-0002668.
PR 19-APR-1996; 96GB-0008164.
XX
PA (IMUT-) IMUTRAN LTD.
PA (QONE-) Q-ONE BIOTECH LTD.
XX
PI Galbraith DN, Haworth C, Lees GM, Smith KT;
XX
DR WPI; 1997-535851/49.
XX
PT Polynucleotide encoding porcine retrovirus expression product -
PT useful to develop products for use in vaccines, diagnosis and
PT xeno-transplantation

XX PS Claim 4; Fig 2; 69pp; English.

XX CC This DNA sequence encodes the porcine retrovirus (POEV) virion core

CC CC polypeptide (GAG), polymerase (POL) and envelope (ENV) proteins. These

CC CC proteins can be used to develop viral vaccines, antisense nucleic acids,

CC CC ribozymes and other antiviral agents. They can also be used in

CC CC xeno-transplantation technology and as diagnostic tools.

XX XX

SO Sequence 8196 BP; 2165 A; 2061 C; 2147 G; 1820 T; 3 other;

Query Match 100.0%; Score 20; DB 18; Length 8196;

Best Local Similarity 100.0%; Pred. No. 2.4;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gatgctctctctgccccttg 20

Db 5226 gatgctctctgccccttg 5245

RESULT 6

AAV09700 AAV09700 standard; DNA; 8209 BP.

XX AC AAV09700;

XX DT 20-MAY-1998 (first entry)

XX DE Porcine retrovirus DNA encoding the LTR and GAG, POL and ENV proteins.

XX XX Porcine retrovirus; POEV; POL protein; ENV protein; GAG protein;

KW virion core polypeptide; polymerase protein; envelope protein;

KW vaccine; diagnosis; xenotransplantation; prophylactic; therapeutic; ds.

XX XX Porcine retrovirus.

XX OS

XX FH Key Location/Qualifiers

FT misc_feature 1..61

FT /tag= a

FT /note= "R LTR domain"

FT 1..588

FT /tag= b

FT 62..143

FT /tag= c

FT /note= "U5 LTR domain"

FT 588..2162

FT /tag= d

FT /product= GAG protein

FT /note= "viral core polypeptide"

FT 2163..5747

FT /tag= e

FT /product= POL protein

FT /note= "polymerase peptide"

FT 5620..7590

FT /tag= f

FT /product= ENV protein

FT /note= "envelope protein"

FT 7591..8209

FT /tag= g

FT 7638..8106

FT /tag= h

FT /note= "U3 LTR domain"

FT 7643..7686

FT /tag= i

FT /note= "OCT-1, c-Myp, LfF-1, E47 and ETS-1"

FT 7719..7739

FT /tag= j

FT /note= "Binding site for transcription factors

FT 7744..7754

FT /tag= k

FT protein_bind

FT FT protein_bind

FT /note= "Binding site for transcription factor AP-1/TR"

FT 7773..7785

FT /tag= l

FT /note= "Binding site for transcription factor ETS-1/GATA"

FT 7793..7834

FT /tag= m

FT /note= "Binding site for transcription factors

FT 7841..7851

FT /tag= n

FT /note= "Binding site for transcription factor AP-1"

FT 7873..7883

FT /tag= o

FT /note= "Binding site for transcription factor AP-1"

FT 7885..7914

FT /tag= p

FT /note= "Binding site for transcription factors

FT 7928..7953

FT /tag= q

FT /note= "Binding site for transcription factors E47

FT 7964..7968

FT /tag= r

FT 7985..7997

FT /tag= s

FT /note= "Binding site for transcription factor GATA"

FT 8003..8007

FT /tag= t

FT 8024..8036

FT /tag= u

FT /note= "Binding site for transcription factor GATA"

FT 8042..8046

FT /tag= v

FT 8049..8060

FT /tag= w

FT /note= "Binding site for transcription factor

FT 8075..8090

FT /tag= x

FT 1..61

FT /tag= y

FT /note= "R LTR domain"

FT 8127..8135

FT /tag= z

FT /note= "PolyA downstream element"

FT 8161..8166

FT /tag= aa

XX PN WO9740167-A1.

XX PD 30-OCT-1997.

XX XX

XX PF 18-APR-1997; 97WO-GB01087.

XX PR 10-FEB-1997; 97GB-0002668.

XX PR 19-APR-1996; 96GB-0008164.

XX PA (IMDT-) IMOTRAN LTD.

XX PA (COMET-) Q-ONE BIOTECH LTD.

XX PI Galbraith DN, Haworth C, Lees GM, Smith KT;

XX WFI; 1997-535851/49.

DR P-PDB; AAM39271; AAM39272; AAM39273.

XX

PT Polynucleotide encoding porcine retrovirus expression product -

PT useful to develop products for use in vaccines, diagnosis and

PT xeno-transplantation

XX PS Claim 4; Fig 3; 69pp; English.

XX XX This DNA sequence encodes the porcine retrovirus (POEV) virion core

XX XX Xenotransplantation; infectious agent; vaccine; ds.
KW XX Porcine endogenous retrovirus.
XX OS WO200071726-A1.
XX PD 30-NOV-2000.
XX PF 24-MAY-2000; 2000WO-US14296.
XX PR 24-MAY-1999; 99US-0135631.
XX PA (MAYO-) MAYO MEDICAL VENTURES.
XX PI Federspiel MJ;
XX DR WPI: 2001-032041/04.
XX PT Inhibiting or preventing infectious agent transmission in mammalian
XX PT transplant recipients, by introducing recombinant DNA comprising DNA
XX PT encoding extracellular proteins of the agent into donor cells, such as
XX PT swine cells -
XX PS Claim 16; Page 109-111; 144pp; English.
XX CC The present invention provides a method to prevent the transmission of
XX CC infectious agents during xenotransplantation. This involves introducing
XX CC to donor swine cells a recombinant DNA encoding a peptide fragment from
XX CC the infectious agent, and then introducing these cells into the
XX CC transplant recipient.
SQ Sequence 4918 BP; 1382 A; 1173 C; 1183 G; 1180 T; 0 other;

Query Match 92.0%; Score 18.4; DB 22; Length 4918;
Best Local Similarity 95.0%; Pred. No. 13;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 gatgctctctgccccttg 20
||| ||||| ||||| |||||
DB 1804 gatagctctctgccccttg 1823

RESULT 10
AAC67021
ID AAC67021 standard; DNA: 6076 BP.
XX AC AAC67021;
XX DT 27-MAR-2001 (first entry)
XX DE PERV env protein coding sequence SEQ ID NO: 21.
XX KW Xenotransplantation; infectious agent; vaccine; ds.
XX OS Porcine endogenous retrovirus.
XX OS WO200071726-A1.
XX PN 30-NOV-2000.
XX PD 24-MAY-2000; 2000WO-US14296.
XX PF 24-MAY-1999; 99US-0135631.
XX PR (MAYO-) MAYO MEDICAL VENTURES.
XX PA Federspiel MJ;
XX PI WPI: 2001-032041/04.
XX DR Inhibiting or preventing infectious agent transmission in mammalian
XX PT

PT transplant recipients, by introducing recombinant DNA comprising DNA
PT encoding extracellular proteins of the agent into donor cells, such as
PT swine cells -
XX PS Claim 16; Page 107-109; 144pp; English.
XX CC The present invention provides a method to prevent the transmission of
XX CC infectious agents during xenotransplantation. This involves introducing
XX CC to donor swine cells a recombinant DNA encoding a peptide fragment from
XX CC the infectious agent, and then introducing these cells into the
XX CC transplant recipient.
SQ Sequence 6076 BP; 1613 A; 1512 C; 1609 G; 1342 T; 0 other;

Query Match 92.0%; Score 18.4; DB 22; Length 6076;
Best Local Similarity 95.0%; Pred. No. 13;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 gatgctctctgccccttg 20
||| ||||| ||||| |||||
DB 5835 gatagctctctgccccttg 5854

RESULT 11
AAC67032
ID AAC67032 standard; DNA: 6076 BP.
XX AC AAC67032;
XX DT 27-MAR-2001 (first entry)
XX DE PERV env protein coding sequence SEQ ID NO: 32.
XX KW Xenotransplantation; infectious agent; vaccine; ds.
XX OS Porcine endogenous retrovirus.
XX OS WO200071726-A1.
XX PN 30-NOV-2000.
XX PD 24-MAY-2000; 2000WO-US14296.
XX PF 24-MAY-1999; 99US-0135631.
XX PR (MAYO-) MAYO MEDICAL VENTURES.
XX PA Federspiel MJ;
XX PI WPI: 2001-032041/04.
XX DR Inhibiting or preventing infectious agent transmission in mammalian
XX PT transplant recipients, by introducing recombinant DNA comprising DNA
XX PT encoding extracellular proteins of the agent into donor cells, such as
XX PT swine cells -
XX PS Claim 16; Page 117-119; 144pp; English.
XX CC The present invention provides a method to prevent the transmission of
XX CC infectious agents during xenotransplantation. This involves introducing
XX CC to donor swine cells a recombinant DNA encoding a peptide fragment from
XX CC the infectious agent, and then introducing these cells into the
XX CC transplant recipient.
SQ Sequence 6076 BP; 1619 A; 1505 C; 1598 G; 1354 T; 0 other;

Query Match 92.0%; Score 18.4; DB 22; Length 6076;
Best Local Similarity 95.0%; Pred. No. 13;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 gatgctctctgccccttg 20

Db 5835 gatagctctcctgaccttg 5854

RESULT 12

AAC67019 ID AAC67019 standard; DNA: 7362 BP.

AC AAC67019;

DE 27-MAR-2001 (first entry)

DE PERV env protein coding sequence SEQ ID NO: 19.

KW Xenotransplantation; infectious agent; vaccine; ds.

OS Porcine endogenous retrovirus.

XX WO200071726-A1.

PD 30-NOV-2000.

PF 24-MAY-2000; 2000MO-US14296.

PR 24-MAY-1999; 99US-0135631.

PA (MAYO-) MAYO MEDICAL VENTURES.

PI Federspiel MJ;

DR WPI; 2001-032041/04.

PT Inhibiting or preventing infectious agent transmission in mammalian

PT transplant recipients, by introducing recombinant DNA comprising DNA

PT encoding extracellular proteins of the agent into donor cells, such as

XX swine cells -

PS Claim 16; Page 101-104; 144pp; English.

CC The present invention provides a method to prevent the transmission of

CC infectious agents during xenotransplantation. This involves introducing

CC to donor swine cells a recombinant DNA encoding a peptide fragment from

CC the infectious agent, and then introducing these cells into the

CC transplant recipient.

SQ Sequence 7362 BP; 1997 A; 1821 C; 1881 G; 1663 T; 0 other;

Query Match 92.0%; Score 18.4; DB 22; Length 7362;

Best Local Similarity 95.0%; Pred. No. 13;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 gatagctctcctgaccttg 20

5900 gatagctctcctgaccttg 5919

RESULT 13

AAC67023 ID AAC67023 standard; DNA: 7873 BP.

AC AAC67023;

DE 27-MAR-2001 (first entry)

DE PERV env protein coding sequence SEQ ID NO: 23.

KW Xenotransplantation; infectious agent; vaccine; ds.

OS Porcine endogenous retrovirus.

XX WO200071726-A1.

PD 30-NOV-2000.

PF 24-MAY-2000; 2000MO-US14296.

PR 24-MAY-1999; 99US-0135631.

PA (MAYO-) MAYO MEDICAL VENTURES.

PI Federspiel MJ;

DR WPI; 2001-032041/04.

PT Inhibiting or preventing infectious agent transmission in mammalian

PT transplant recipients, by introducing recombinant DNA comprising DNA

PT encoding extracellular proteins of the agent into donor cells, such as

XX swine cells -

PS Claim 16; Page 112-115; 144pp; English.

CC The present invention provides a method to prevent the transmission of

CC infectious agents during xenotransplantation. This involves introducing

CC to donor swine cells a recombinant DNA encoding a peptide fragment from

CC the infectious agent, and then introducing these cells into the

CC transplant recipient.

SQ Sequence 7873 BP; 2200 A; 1913 C; 1962 G; 1798 T; 0 other;

Query Match 92.0%; Score 18.4; DB 22; Length 7873;

Best Local Similarity 95.0%; Pred. No. 13;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 gatagctctcctgaccttg 20

4629 gatagctctcctgaccttg 4648

RESULT 14

AAT74884 ID AAT74884 standard; CDNA: 7892 BP.

AC AAT74884;

DE 09-FEB-1998 (first entry)

DE Miniature swine retrovirus CDNA.

KW Retrovirus; porcine; GAG protein; POL protein; ENV protein;

KW xenotransplantation; infectious; provirus; organ transplant; donor;

KW activated virus; PCR; ss.

OS Porcine retrovirus.

XX Key

XX CDS

XX mat_peptide

XX CDS

XX mat_peptide

XX CDS

XX mat_peptide

XX CDS

WO9721836-A1.

19-JUN-1997.


```

XX 13-DEC-1996; 96WO-US19680.
XX 14-DEC-1995; 95US-0572645.
XX (GEHO) GEN HOSPITAL CORP.
XX Fishman JA;
XX WPI: 1997-332804/30.
XX P-PSDB: AAM32096-W32098.
XX New nucleic acid from porcine retroviruses - used for detecting
XX viruses in transplant or other tissue and for assessing risk of
XX transmitting infection to graft recipient
XX Claim 22; Fig 3; 128bp; English.
XX This CDNA sequence represents a porcine retrovirus from miniature swine
XX containing the putative coding regions for viral GAG, POL and ENV
XX proteins. This sequence and PCR fragments generated from it
XX (see AAT74812-T74882) can be used to screen organs for the presence of
XX porcine retroviruses prior to xenotransplantation. Transplantation can
XX increase the likelihood of retroviral activation if intact and
XX infectious proviruses are present. The porcine retroviral sequence can be
XX used to generate probes to determine the level (e.g. copy number) of
XX intact (i.e. potentially replicating) porcine provirus sequences in a
XX strain of xenograft transplantation donors. It can be used to detect
XX mutations, genetic lesions or viral recombinants and also to determine
XX the histological localisation of activated retroviruses. Using Polymerase
XX Chain Reaction DNA Quantitation (PDQ) on blood mononuclear cells,
XX infectivity titration and susceptibility testing can be performed.
XX Ultimately animal donors without intact porcine retroviral sequences or a
XX lower copy number of viral elements could be selected.
XX Sequence 7892 BP; 2191 A; 1915 C; 1980 G; 1806 T; 0 other;
SQ
Query Match 92.0%; Score 18.4; DB 18; Length 7892;
Best Local Similarity 95.0%; Pred. No. 13;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 gatgctctcctgccttgg 20
Db 5237 gatagctctcctgccttgg 5256
RESULT 15
AAT74811
ID AAT74811 standard; CDNA; 8060 BP.
XX
XX AAT74811;
XX
XX 11-FEB-1998 (first entry)
XX Porcine retrovirus Tsukuba-1 CDNA.
XX
XX Porcine retrovirus Tsukuba-1 CDNA.
XX
XX Retrovirus; porcine; GAG protein; POL protein; ENV protein;
XX xenotransplantation; infectious; provirus; organ transplant; donor;
XX activated virus; Tsukuba-1; PCR; ss.
XX
XX Porcine retrovirus.
XX
XX Key Location/Qualifiers
XX CDS 86..2002 /*tag= a
XX /*note= "putative ENV protein"
XX CDS 3112..4686 /*tag= b
XX /*note= "putative GAG protein"
XX CDS 4871..8060 /*tag= c
XX /*note= "putative POL protein (partial)"
XX

```

```

XX WO9721836-A1.
XX 19-JUN-1997.
XX
XX 13-DEC-1996; 96WO-US19680.
XX 14-DEC-1995; 95US-0572645.
XX (GEHO) GEN HOSPITAL CORP.
XX Fishman JA;
XX WPI: 1997-332804/30.
XX New nucleic acid from porcine retroviruses - used for detecting
XX viruses in transplant or other tissue and for assessing risk of
XX transmitting infection to graft recipient
XX Claim 1; Fig 1; 128bp; English.
XX This sequence represents the purified porcine retroviral CDNA
XX sequence of Tsukuba-1 and contains the putative coding regions for viral
XX proteins GAG, POL and ENV. This sequence and PCR fragments generated
XX from the sequence (see AAT74812-T74882) could be used to screen organs
XX for porcine retroviruses prior to xenotransplantation. Transplantation
XX can increase the likelihood of retroviral activation if intact and
XX infectious proviruses are present. The porcine retroviral sequence can be
XX used to generate probes to determine the level (e.g. copy number) of
XX intact (i.e. potentially replicating) porcine provirus sequences in a
XX strain of xenograft transplantation donors. It can be used to detect
XX mutations, genetic lesions or viral recombinants and to determine the
XX histological localisation of activated retroviruses. Using Polymerase Chain
XX Reaction DNA Quantitation (PDQ) on blood mononuclear cells, infectivity
XX titration and susceptibility testing can be performed. Ultimately animal
XX donors without intact porcine retroviral sequences or with a lower copy
XX number of viral elements could be selected.
XX Sequence 8060 BP; 2233 A; 1959 C; 2012 G; 1856 T; 0 other;
SQ
Query Match 92.0%; Score 18.4; DB 18; Length 8060;
Best Local Similarity 95.0%; Pred. No. 13;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 gatgctctcctgccttgg 20
Db 7763 gatagctctcctgccttgg 7782

```

Search completed: February 24, 2002, 03:05:29
Job time: 19495 sec

